Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

Psychological Research

JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

EDITED BY
CARL MURCHISON

JOHN E. ANDERSON
University of Minnesota

S. K. BOSE
University of Calcutta

CHARLOTTE BÜHLER
University of Southern California

CYRIL BURT
University of London

LEONARD CARMICHAEL
Smithsonian Institution

SANTE DE SANCTIS
R. Università di Roma

ARNOLD GESELL
Yale University

WILLIAM HEALY
Judge Baker Foundation, Boston

ARTHUR T. JERSILD
Teachers College, Columbia University

BUFORD JOHNSON
The Johns Hopkins University

HAROLD E. JONES
University of California

TRUMAN L. KELLEY
Harvard University

H. P. MAITI
Institute of Child Development, Ahmedabad, India

TORAO OBONAI
Tokyo University, Japan

YOSHIKAZU OHWAKI
Tohoku University, Japan

HENRI PIÉRON
Université de Paris

ROBERT R. SEARS
Stanford University

GEORGE D. STODDARD
University of Illinois

C. J. WARDEN
Columbia University

JOHN B. WATSON
New York City

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold under the restrictions of a bilateral contract, and may not be resold for a period of five years from date of publication.

VOLUME 94

1959

& Psyl. Research
9.3.81
A.M
~~J745~~ J745

Copyright, 1959, by The Journal Press

Published Quarterly by The Journal Press
Provincetown, Massachusetts, U. S. A.

AUTHOR INDEX

Aarons, Louis 101
Albert, Robert S. 287
Blair, Margaret A. 93
Brown, W. Lynn 107, 211
Carr, Richard M. 107
Cassel, Russell N. 161
Church, Joseph 169
Cooper, Joseph B. 93
Coslett, Stephen B. 215
French, Joseph L. 131
Goertzen, Stanley W. 29
Goss, Albert E. 221
Haggerty, Arthur D. 205
Harms, Irene E. 3
Henderson, Harold L. 143
James, W. T. 199
Kaden, Stanley 267
Lacey, Harvey 221
Levine, Seymour 77
Maslow, A. H. 43
Masten, Sherman H. 143
McDowell, Arnold A. 35, 211
Medinnus, Gene R. 253
Nissen, Henry W. 35
Phillips, Leslie 267
Reed, G. F. 67
Richardson, Claudia 169
Riesen, Austin H. 101
Rychlak, Joseph H. 153
Shaklee, Alfred B. 215
Simpson, James E. 263
Smith, A. C. 67
Spiker, Charles C. 3
Stott, D. H. 233
Symmes, David 181
Trent, Richard D. 85
Van Krevelen, Alice 23
Waldman, Marvin 267
Zuk, G. H. 113

TABLE OF CONTENTS

Factors associated with the performance of young children on intelligence scales and tests of speech development 3
By Irene E. Harms and Charles C. Spiker

Stability of adolescents' judgments of an ambiguous figure after knowledge of others' judgments of same figure 23
By Alice Van Krevelen

Factors relating to opinions of seventh grade children regarding the acceptability of certain behaviors in the peer group 29
By Stanley M. Goertzen

Solution of a bi-manual coördination problem by monkeys and chimpanzees . 35
By Arnold A. McDowell and Henry W. Nissen

Cognition of being in the peak experiences 43
By A. H. Maslow

An experimental investigation of the relative speeds of left and right-handed writers 67
By A. C. Smith and G. F. Reed

Emotionality and aggressive behavior in the mouse as a function of infantile experience 77
By Seymour Levine

Anxiety and accuracy of perception of sociometric status among institutionalized delinquent boys 85
By Richard D. Trent

Parent evaluation as a determiner of ideology 93
By Joseph B. Cooper and Margaret A. Blair

Apparatus for studying visual discriminations of movement in small mammals 101
By Louis Aarons and Austin H. Riesen

The effect of the introduction of novel stimuli upon manipulation in rhesus monkeys 107
By Richard M. Carr and W. Lynn Brown

The mind as an optical system: A study of perceptual reproduction in the normal and mentally retarded 113
By G. H. Zuk

Intellectual appraisal of physically handicapped children 131
By Joseph L. French

Six predictors of college achievement 143
By Harold L. Henderson and Sherman H. Masten

BOOKS 147

BOOKS RECENTLY PUBLISHED 147

Self-confidence, ability, and the interest-value of tasks 153
By Joseph H. Rychlak

Comparing the effectiveness of the ego-strength Q-sort test by use of R- and Q-methodologies 161
By Russell N. Cassel

A developmental analysis of proverb interpretations 169
By Claudia Richardson and Joseph Church

Anxiety reduction and novelty as goals of visual exploration by monkeys . 181
By David Symmes

Behavior of the opossum in the Fink Arrow Maze 199
By W. T. James

The effects of long-term hospitalization or institutionalization upon the language development of children 205
By Arthur D. Haggerty

The relation between pre-irradiation volitional activity and post-irradiation death order in the rat 211
By Arnold A. McDowell and W. Lynn Brown

Parallel observations on avoidance behavior and endocrine response under reserpine 215
By Stephen B. Coslett and Alfred B. Shaklee

Conceptual block sorting as a function of number, pattern of assignment, and strength of labeling responses 221
By Harvey Lacey and Albert E. Goss

Infantile illness and subsequent mental and emotional development . . . 233
By D. H. Stott

Immanent justice in children: A review of the literature and additional data 253
By Gene R. Medinnus

The influence of auditory stimulation on aniseikonic perception: II and III. Business school women and college women 263
By James E. Simpson

Rorschach indices of developmental level 267
By Leslie Phillips, Stanley Kaden, and Marvin Waldman

The function of verbal labels in the discrimination of subtle stimulus differences 287
By Robert S. Albert

BOOKS 297

BOOKS RECENTLY PUBLISHED 297

The Journal of Genetic Psychology, 1959, **94**, 3-22.

FACTORS ASSOCIATED WITH THE PERFORMANCE OF YOUNG CHILDREN ON INTELLIGENCE SCALES AND TESTS OF SPEECH DEVELOPMENT*

Iowa Child Welfare Research Station

IRENE E. HARMS AND CHARLES C. SPIKER[1]

A. INTRODUCTION

This study was designed to investigate the consistency of performance of infants and young children on two intelligence scales and on certain indices of speech development, the relationship between the two intelligence scales, and the relationships among the intelligence tests and the speech indices.

The most common usage of the results of intelligence tests involves the assumption that the score of a given subject does not fluctuate markedly from day to day. That is, a test is not ordinarily given with the purpose of evaluating a child's level of performance on a single day, for a given week, nor even for a month. It is, on the contrary, generally expected that the results of a test given today will have predictive value for that child at some later time. To avoid gratuitous assumptions, this assumed consistency of behavior should be studied. It is apparent that the method of "split-half reliability" does not necessarily provide a valid estimate of the day-to-day, week-to-week fluctuations in level of performance which may occur. Yet this type of approach to the consistency of performance is often made by test constructors. In particular, Cattell (2) reports split-half reliabilities for seven age levels from 3 to 30 months. Ku[illegible]mann (15, 16) very largely ignores the problem, pointing out that determining the reliability of a test score is a psychological job, not something which can be accomplished mathematically. This point of view appears to recognize and properly categorize the difficulties involved in discovering the factors which would permit the prediction of the test scores of individual children, but it does not emphasize the usefulness for clinical purposes of some estimate of the average fluctuation of test scores.

---

*Received in the Editorial Office early in 1954, destroyed in the fire of May 1, 1956, and now rewritten and published as of March 21, 1958.

[1] This article is based on separate dissertations by the writers submitted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in the Child Welfare Research Station, State University of Iowa. The writers express their appreciation to Dr. Orvis C. Irwin under whose direction the research was performed.

Few studies of the consistency of performance on intelligence tests over short time intervals contribute anything to this problem for children under 2½ years of age. Summaries of these studies are presented by Goodenough and Maurer (5) and Maurer (18). The results of the studies suggest that under carefully controlled conditions, groups of children perform consistently over short intervals of time on the same tests and on different equated forms of tests, but that the degree of consistency is considerably less when entirely different scales of intelligence are used. Further investigation in this area would seem promising.

Considerable progress has been made in the study of spontaneous vocalization of infants and young children with the development of a behavioral unit of sampling as opposed to a time sampling technique (14). The unit in this case is the breath-unit—the sounds produced by the subject on the exhalation phase of the breath-cycle. Using this technique and a constant number of units per record, several indices of speech development have been devised and the relationship of these indices to other important developmental variables have been studied. Among others, the indices of consonant, vowel, and phoneme type, and consonant, vowel, and phoneme frequency have been devised, and studies made of the relationships of these to such variables and conditions as chronological age (8, 12, 13), sex of the child (8, 12), paternal occupation, (9) feeblemindedness (7), institutionalization (1), and intelligence (11). Many of the results of these studies have proved to be both interesting and enlightening, and further study appears profitable.

One important aspect of the method described above which has not yet received sufficient attention is the consistency of the behavior of individual subjects on the speech indices from one sample of speech to another. That is, do subjects tend to maintain the same relative scores on the indices from one occasion to the next? The answer to this question should indicate whether or not it would be useful to devise standardized situations in which to take the speech samples.

B. Statement of the Problem

The study reported here was designed to attempt to answer each of the following questions concerning the performance of children between the ages of 16 and 30 months on the intelligence tests and the indices of spontaneous speech sounds:

1. Is the degree of consistency of performance on the intelligence scales dependent on such factors as examiners, scale order, recent previous experience with the scale, and the *CA* and sex of the children?

2. How consistently do the children perform with respect to the two speech indices—phoneme type and phoneme frequency—from a first speech sample to a second two to seven days later?

3. What is the relationship between the two intelligence tests under the conditions set up in the present study?

4. What are the relationships among the speech indices and the intelligence scales?

5. Are the *IQ* means for groups of children related to the scale used, the examiners administering the test, the order of giving the test, and the recent previous experience of the child with the test?

6. How do the speech indices relate to the age and sex of the subjects?

C. EXPERIMENTAL PROCEDURE

1. *Subjects*

The subjects used in this study were 80 children ranging in *CA* from 16 through 29 months, living with white, English-speaking parents, and having no obvious physical defects. There were 10 subjects in each of eight age-sex subgroups, the *CA* means for which are given in Table 1. The parents of all children were either students or faculty members at the State University of Iowa. Seventy-eight of the 80 subjects lived in a housing area maintained by the University for married students and staff members. The parents of two subjects lived in off-campus housing units. Two additional subjects were eliminated from the sample because it was not possible to obtain complete records of non-crying sounds nor to administer the intelligence tests satisfactorily.

TABLE 1
THE *CA* MEANS AND RANGES FOR THE AGE AND SEX SUBGROUPS OF 10 CASES EACH

	Age range			Boys	Girls
Group	Mos. Days		Mos. Days	Mean	Mean
1	16 - 1	to	17 - 30	16.81	17.09
2	20 - 1	to	21 - 30	21.29	21.18
3	24 - 1	to	25 - 30	25.09	25.20
4	28 - 1	to	29 - 30	28.97	28.84

2. *General Experimental Design*

The design of the experiment is best described with reference to the two sessions conducted with each child. In the first session, the Kuhlmann (16) and Cattell (2) intelligence scales were successively administered, and a 60 breath-unit sample of the child's vocalization was obtained. In the second session, the same procedure was carried out again. For no child was the second session separated from the first by more than 7 days nor by less than 36 hours, and for the great majority of children, the interval was 2, 3, and 4 days.

Since it was desired to study the effects of the order of administration of the intelligence tests and examiner differences upon the test-retest correlations, it was necessary to vary these conditions within each age-sex subgroup. The basic pattern for each age-sex subgroup, and for the total group, is given in Table 2, with the number of cases in each condition. In this table, *C* represents Cattell; *K*, Kuhlmann; *H*, Examiner *H*; *S*, Examiner *S*; and the order of the test administration is indicated by the position of the letters. Thus, *CK:HS* indicates that the Cattell was given first and the Kuhlmann was given second during both sessions, and that Examiner *H* administered these tests during the first session, Examiner *S* during the second. In any comparison of age-groups or of sex-groups, the order of administering the tests and the examiner differences are thus counterbalanced. Each subject had the same test order (*CK* or *KC*) for both the first and second sessions.

TABLE 2
THE NUMBER OF CASES IN EACH SCALE ORDER AND EXAMINER CONDITION FOR EACH AGE-SEX SUBGROUP AND FOR THE TOTAL GROUP

Condition	Number per age-sex subgroup	Number total group
CK:HS	1	8
CK:SH	1	8
CK:HH	3	24
KC:HS	1	8
KC:SH	1	8
KC:HH	3	24
Total	10	80

The speech samples for all children in both sessions were recorded by one examiner (*S*). These samples were of the subject's spontaneous speech; i.e., the subject's verbal responses to questions by the examiner were never recorded. In general, when Examiner *H* administered the intelligence tests, the subject's speech was recorded during the administration of the test. When Examiner *S* administered the tests, the speech record was ordinarily obtained following the administration of the tests. This systematic difference is also counterbalanced in any comparison of age-groups or sex-groups.

With the exception of two subjects, an adult (mother, father, grandmother, or familiar adult) remained in the administration room during both sessions of testing the child. The adult was not permitted to interfere with the administration of the test nor to coach the child in his responses.

3. *Specific Procedures*

One of the major problems in studies of this kind concerns the objectivity (observer agreement) of the events studied. This is of considerable importance in determining the correctness of responses to intelligence test items; it is of even more importance in the recording of speech sounds. Except where the test manuals were ambiguous, tests were administered and scored in accordance with the authors' instructions. To obtain a maximum of uniformity, preliminary tests were administered by each examiner and observed by the other. Agreement was thus reached on standards for administration and scoring and for the substitution of items. Additional conferences were held during the early stages of the main study when other doubtful items were noted. Examiner differences in test scores should be minimized by this procedure and should not be regarded as representative of what might be found under other conditions.

Within each session, an attempt was made to complete the first test before proceeding with the second. When the Cattell was given first, it was always completed before items from the Kuhlmann were given. In some instances when the Kuhlmann test was given first, it was necessary to begin the Cattell before the Kuhlmann was completed because subjects lost interest in the Kuhlmann items. The records for these subjects were not analyzed separately.

Unfortunately, it is not possible to give a simple, exhaustive, and practical description of the objectivity of the recording of the subjects' speech sounds. Part of this problem has been discussed in detail by Irwin and Chen (10), and still other aspects of it by Spiker (20). However, the difficulties involved do not preclude descriptions of observer agreement which are useful for certain purposes.

Examiner *S* was trained in speech recording by an observer[2] whose competence in recording had been previously demonstrated (14). The latter had taken all courses in phonetics offered at the university, and his original training was conducted by a trained phonetician. The training of Examiner *S* was carried out over a period of a year, using as subjects children from 9 to 30 months of age. After the initial training period, Examiner *S* and the previously trained observer conducted a preliminary study in order to evaluate the former's proficiency in recording. Fifteen subjects were selected, ranging in age from 14 through 29 months, and 30 breath-unit samples were obtained on each subject by each observer under the following conditions:

---

[2] The writers wish to express gratitude to Dr. Orvis C. Irwin for his liberal contribution of time and guidance to the training of the junior author in the recording of speech sounds.

1. Both observers visited each subject in his own home and recorded simultaneously, but independently, the sounds produced on the exhalation phase of selected breath cycles.

2. These sounds were recorded in the International Phonetic Alphabet, after Fairbanks (4).

3. For each subject, the observers' records were analyzed by transferring the data from the records to a tally sheet, thereby indicating the frequency with which each observer recorded each sound.

Two methods were then used to obtain numerical descriptions of the degree of observer agreement. The first method yields a percentage agreement index. Within the tally sheet for each subject, a percentage agreement index was obtained for each phoneme by means of the formula:

$$PA_p = 100 \frac{f_L}{f_G} \tag{1}$$

where PA_p is the percentage agreement for a given phoneme for a given subject, f_L is the lesser, and f_G is the greater of the two frequencies reported by the two observers for that phoneme. For each individual, an average agreement index was computed which was based on the indices for each phoneme separately, and was obtained by means of the following:

$$PA_i = \frac{\Sigma[(f_L + f_G)(PA_p)]}{\Sigma(f_L + f_G)} \tag{2}$$

where the summation is over all phonemes for the individual and the remaining terms are the same as in equation (1). The average percentage agreement for the total sample was obtained by:

$$PA_s = \frac{\Sigma[PA_i \ \Sigma(f_L + f_G)]}{\Sigma\Sigma(f_L + f_G)} \tag{3}$$

where the second summation is over all subjects and the other terms have the same meanings as in Equations 1 and 2. It is apparent that this method utilizes weights which serve to reduce the contribution to the average percentage agreement of those phonemes which occur infrequently (or never) in the records of the two observers. The average percentage agreement for the total sample by this method was 84 per cent, with a range from 62 per cent to 96 per cent for the 15 individual cases.

The second method of computing observer agreement makes use of correlational techniques. For this purpose, the two indices of speech sounds of primary concern in this study were utilized—the phoneme type index and the phoneme frequency index, *PT* and *PF*, respectively. The definitions of these concepts appear below:

PT. The total number of *different* sounds in the record of the observer for a given subject.

PF. The total number of sounds, *disregarding type*, in the record of the observer for a given subject.

For each of the 15 subjects in the observer-agreement sample, these indices were determined from the 30 breath-unit records of each of the two observers. Product moment correlation coefficients were then computed between corresponding indices obtained from the records of the two observers. The values obtained were .88 and .95 for the *PT* and *PF* indices, respectively.

The procedure for recording speech sounds during the major study did not differ greatly from that used in obtaining data for the study of observer-agreement. The major difference is that samples of 60 instead of 30 breath-units were recorded in each session. Irwin and Chen (10) found that the same results can be obtained by recording sounds produced on every second or fourth breath as can be obtained by recording sounds on successive breaths. Hence, only those vocalizations of the subjects were recorded which were not made in response to verbal questioning, which did not occur during the observer's recording of previously produced sounds, and which were clearly audible to the observer.

The times required to complete single sessions ranged from approximately 45 to 90 minutes.

D. RESULTS

1. *The Consistency of Performance*

a. The effects of different examiners on consistency of performance on the intelligence scales. The test-retest correlations of the two intelligence scales provide an indication of the consistency of performance of the subjects. A comparison of the test-retest correlations for subjects tested by the same examiner and those tested by different examiners in the two sessions yields an estimate of the influence of examiners upon the consistency of performance of the subjects. Examiner *H* administered the scales to 48 subjects in both sessions. Examiner *S* administered the scales to 16 subjects during the first session and to another 16 during the second session, while Examiner *H* administered the scales to these same subjects in the reverse order. Which examiner gave the scales in which session, and the order of administering the scales, were therefore counterbalanced for these 32 subjects (Table 2). The test-retest correlation for the 48 subjects given both test and retest by the same examiner was .81 for the Kuhlmann, and for the 32 subjects given the first test by one examiner and the second by the other examiner .75. The difference between these two values was smaller than the standard error, suggesting that under the conditions of this experiment, the consistency of performance is not dependent upon a difference in examiners. When the Cattell results were analyzed in the same way, the corresponding values were .82 for the 48 subjects tested each time by the same examiner and .93 for those tested by different examiners in the two sessions. This latter difference was significant beyond the 5 per cent level, but it should be noted that it was in the opposite direction from what would be expected and from the results obtained for the Kuhlmann. It therefore seems reasonable to assume that under the conditions of the present study, a change in examiners does not affect test-retest reliability.

b. The effects of age and sex of subjects on consistency of performance on intelligence scales. When the total group was broken down into age groups and the Kuhlmann test-retest correlations obtained within each age level, the values were .79, .80, .80, and .85 for Groups 1, 2, 3, and 4, respectively. The hypothesis that these four sample values arose from a common bivariate population was tested according to a procedure set forth by Snedecor (19) and no evidence was found for the rejection of the hypothesis. On the assumption of the correctness of the hypothesis, an estimate of .81 was obtained for the population value. This value may be compared with the .79 which was obtained from correlating the entire 80 cases in the total group, a value which was also obtained for each of the two sex groups of 40 cases each. Hence, there is no evidence for differential consistency of performance on the Kuhlmann for groups differing in either age or sex.

The same analysis was made for the Cattell test. The test-retest correlations for age Groups 1, 2, 3, and 4 were .80, .96, .80, and .90, respectively. The same test of the null hypothesis yielded a probability value between .05 and .02, thus calling into question the hypothesis of a common parameter. The value which seemed to produce the significant probability was the .96 for the 20 and 21 month group. The writers have no plausible explanation as to why there was a significantly higher value for this group, and are inclined to attribute this deviation to chance. On the assumption of a common parameter for the four age groups, the average Cattell test-retest correlation was .89. This value may be compared with the .87 which was obtained for the entire 80 cases, disregarding age and sex, and to the .87 and .85 for the 40 boys and 40 girls respectively, where age has been ignored. Hence, while there is no evidence for differential consistency of performance for groups differing in sex, there is some question as to whether or not groups of subjects differing in age within the ranges studied here, would perform with equal consistency on the Cattell scale.

c. The effects of the order of administering the scales on the consistency of performance on the intelligence scales. One-half the total group received the Kuhlmann first in both sessions and the other half received the Cattell first. Age, sex, and examiner differences are counterbalanced for these two subgroups. For a given scale, it is conceivable that the consistency of performance may be dependent upon the order of presenting the scale. This question can be answered for the Kuhlmann test by comparing the consistency of performance of the 40 subjects given the Kuhlmann first with that for the 40 given the Kuhlmann after having been given the Cattell. The test-retest correlation for the former was .68 and for the latter, .89. The difference between these two values was significant at the 1 per cent level of confidence. Corresponding values for the test-retest correlations on the Cattell were .82 for those given the Cattell first and .90 for those given the Kuhlmann first. While this difference was not statistically significant, it should be noted that it is consistent with the result obtained for the Kuhlmann in the sense that the performance is more consistent on the Cattell when the scale order provides for a test preceding the Cattell. This evidence suggests that, for a given scale, a scale order which provides for a preliminary test period will result in greater consistency of performance on that scale.

d. The consistency of performance on the speech indices. As indicated in a previous section, a 60 breath-unit record of speech sounds was collected for each subject in each session. From these records, the number of phoneme types (different sounds) and the total number of sounds (ignoring type) were determined. Obviously, these two indices could also be determined for any sub-part of the total 120 breath-units and comparisons could be made among equal sub-parts. In order to determine the stability of the phoneme type index for different sizes of samples of speech sounds, the total 120 breath-units were divided into successive 30 breath-unit records and the numbers of different sounds were counted for the first 30, the first 60, the first 90, and for the entire 120 breath-units. The results of this analysis are presented in Table 3. The type means appear to depend upon the number of breath-units in the samples upon which they were based. The function is a negatively accelerated one, and, judging by the successive gains obtained, a sample greater than 120 breath-units would not be expected to increase the type means appreciably. Since the phoneme type score based on 120 breath-units appears to be more stable than either that based on 60 breath-units or that based on the average for two 60 breath-unit records, it has been used in all analyses except for the study of the consistency of performance.

TABLE 3
THE MEANS OF THE PHONEME TYPE SCORES FOR CUMULATED 30 BREATH-UNIT SAMPLES FOR EACH OF THE FOUR AGE LEVEL GROUPS AND FOR THE TOTAL GROUP

Age group	Number of breath units			
	30	60	90	120
1	14.50	18.10	19.90	21.20
2	19.90	23.50	25.55	26.55
3	23.10	26.10	28.10	28.60
4	24.45	27.75	29.80	30.65
Total	20.49	23.86	25.84	26.75

The consistency of performance on the phoneme type index over the two sessions was studied by correlating the scores on the first session with those on the last session for each of the four age groups. These indices were based on 60 breath-units each, and the relevant correlations are given in Table 4.

TABLE 4
THE TEST-RETEST AND SPLIT-HALF CORRELATIONS FOR EACH OF THE SPEECH INDICES WITHIN FOUR AGE LEVEL GROUPS, WITH AVERAGES FOR TOTAL GROUP

Group	Phoneme type		Phoneme frequency	
	Test-retest	Split-half	Test-retest	Split-half
1	.72	.64	.73	.50
2	.51	.51	.79	.78
3	.71	.78	.57	.68
4	.60	.73	.87	.88
Average	.64	.68	.76	.74
Spearman-Brown estimate		.81		.85

These values did not differ significantly from one another and the best estimate for the entire age range, under the null hypothesis, was .64. By dividing the 60 breath-units for the first session into equal halves—first 30 and second 30—and correlating the two phoneme type scores thus obtained, another estimate of performance consistency was obtained. These "split-half reliabilities" for Groups 1, 2, 3, and 4 were .64, .51, .78, and .73, respectively. Under the null hypothesis, which could readily be retained, the population estimate was .68. When this value was corrected by the Spearman-Brown formula, it was .81.

The same type of analysis was made for the phoneme frequency index and Table 4 presents the results. For both test-retest and split-half, the values obtained at the several age levels did not differ significantly from one to another and appropriate estimates of the common parameters were obtained.

There was little difference between the test-retest and the uncorrected split-half estimates, .76 and .74, but the corrected split-half correlations were higher than the test-retest values for both phoneme type and phoneme frequency. This suggests that the inconsistency of the performance is not due to any fundamental instrument unreliability, but rather to factors which are associated with different situations and with the intervening periods of time.

e. The interrelationships of the intelligence scales and the speech indices. The Relationship Between the Intelligence Scales. The correlation between the Kuhlmann and the Cattell scales gives an indication of the degree to which the tests may be used interchangeably, although not in the sense that the numerical scores may be expected to be equal or roughly equivalent. This relationship has been studied separately for the two sex groups and for the four age groups, and for the first test, the retest, and the average test-retest *IQ*s. There were only minor differences between the sex groups in the degree of correlation on the first test, the retest, and the average of the two tests, and these values are not presented here. The intercorrelations of the two scales for the several age groups are presented in Table 5. The differences among the correlation coefficients obtained for the various age levels were significant for none of the three columns. The best estimate of the assumed common parameter for each column is given in the last line of Table 5. As might be expected with the length of the test increased by combining the two scores, the correlations tended to be slightly higher when the average *IQ*s for the two scales are correlated than for test and retest considered separately. For this reason, the average *IQ* for a subject obtained from his first session and second session scores was used in analyses of the relationships between the intelligence scales and the speech sound indices.

TABLE 5

THE CORRELATIONS BETWEEN THE KUHLMANN AND CATTELL INTELLIGENCE SCALES WITHIN THE FOUR AGE GROUPS, USING THE FIRST TEST, THE RETEST, AND THE TEST-RETEST AVERAGE AS CRITERION SCORES

Age group	First test	Retest	Test-retest average
1	.71	.63	.70
2	.84	.87	.91
3	.80	.78	.82
4	.78	.91	.89
Average	.78	.82	.85

f. The relationship between the speech indices. This relationship was estimated by correlating, within age levels, the phoneme type index based on 120 breath-units with the phoneme frequency index based on an average of the 60 breath-unit records from the two sessions. The resulting values were .65, .57, .55, and .62 for Groups 1, 2, 3, and 4, respectively. There was no evidence against the null hypothesis so that an estimate of .60 was obtained for the entire age range.

g. The relationships between the speech indices and the intelligence scales. These relationships were studied by correlating the average *IQ* of each child on each test for the two sessions with his phoneme frequency scores for the two 60 breath-unit records and his phoneme type score on the 120 breath-unit samples. These correlations were again run within age groups in order to minimize the contribution of *CA* to these relationships. Table 6 presents the resulting correlations for each of the four age groups and the corresponding probabilities that the values in each column would have arisen by chance from random samples from a single bivariate population. While only one of the four probabilities offered strong evidence against the null hypothesis, the consistently low probabilities, together with the three negative correlations in the first age group, raised considerable doubt concerning the homogeneity of the correlations within the columns. For this reason, Group 1 was subsequently omitted and the test for the null hypothesis was made on Groups 2, 3, and 4. The results of this test are presented in Table 7. The

TABLE 6

TABLE OF INTERCORRELATIONS OF TWO SPEECH INDICES AND TWO INTELLIGENCE TESTS FOR THE FOUR AGE GROUPS, WITH RELEVANT PROBABILITIES FOR HAVING ARISEN BY CHANCE FROM A SINGLE BIVARIATE POPULATION

Group	Phoneme type		Phoneme frequency	
	Kuhlmann	Cattell	Kuhlmann	Cattell
1	.25	—.02	—.24	—.26
2	.77	.63	.49	.45
3	.46	.56	.50	.49
4	.64	.71	.39	.29
Probabilities	$.20 > p > .10$	$.05 > p > .02$	$.10 > p > .05$	$.10 > p > .05$

TABLE 7

TABLE OF INTERCORRELATIONS OF TWO SPEECH INDICES AND TWO INTELLIGENCE TESTS FOR AGE GROUPS 2, 3, AND 4, WITH RELEVANT PROBABILITIES FOR HAVING ARISEN BY CHANCE FROM A SINGLE BIVARIATE POPULATION

Group	Phoneme type		Phoneme frequency	
	Kuhlmann	Cattell	Kuhlmann	Cattell
2	.77	.63	.49	.45
3	.46	.56	.50	.49
4	.64	.71	.39	.29
Probabilities	$.50 > p > .30$	$.80 > p > .70$	$.90 > p > .80$	$.80 > p > .70$
Averages	.63	.63	.46	.41

next to the last line of this table shows the relevant probabilities and it may be observed that they are considerably higher and offer no evidence against the null hypothesis. Under this hypothesis, the population estimates were computed and are presented in the last line. The two intelligence scales correlate to about the same degree with each speech index for Groups 2, 3, and 4. Both tests had correlations of .63 with the phoneme type index; the Kuhlmann had a correlation of .45 and the Cattell, .41, with the phoneme frequency index. Each of these four correlations had 51 degrees of freedom (19), and was significant at less than the 1 per cent level of confidence.

E. Group Differences in Performance on Speech and Intelligence Tests

1. *Factors Associated with Group IQ*

a. Examiner differences. Thirty-two children were tested by the two examiners with the age and sex of subjects, the order of administration of the two scales, and the number of tests and retests counterbalanced for each examiner (Table 2). A four-dimensional analysis of variance of the changes in *IQ* from initial to final tests was made, involving scale, scale order, sex, and examiner giving the first test as the classifications. The means for these subgroups of four cases each are given in Table 8 and the summary of the analysis of variance is given in Table 9. The design is a "mixed design" (17), some of the effects requiring intra-individual and some requiring inter-individual variability as the error term for the evaluation of statistical significance. The examiner effect proved to be significant beyond the 5 per cent level of confidence, the gain in *IQ* tending to be significantly greater when Examiner *H* gave the first test and Examiner *S* gave the second. This means that Examiner *S* tended to obtain higher *IQ* scores than did Examiner *H*. No other main effect approached statistical significance. Three first-order

TABLE 8

Mean *IQ* Gains from First to Second Test, Kuhlmann and Cattell, for the Independent Sex-Scale Order Subgroups, Divided According to the Examiner Giving First Test, Ignoring Age Groups, with Four Cases per Cell

		Kuhlmann Examiner first:		Cattell Examiner first:	
Sex	Order	*S*	*H*	*S*	*H*
Boys	*KC*	—1.75	5.75	—2.50	7.00
	CK	—1.25	7.50	0.25	7.75
Girls	*KC*	8.25	10.25	3.75	6.25
	CK	—5.50	1.50	8.00	2.75

and one second-order interactions were significant at less than the 10 per cent level of confidence. These were the sex by scale order, sex by examiners, scales by scale order, and scales by scale order by sex interactions. The relatively high probabilities for these differences would not seem to warrant further discussion until they have been duplicated in other studies. We may, however, conclude that the examiners constitute a significant factor which should be taken into consideration in evaluating the *IQ* of an individual or the *IQ* level of a group of individuals.

TABLE 9

SUMMARY OF THE ANALYSIS OF VARIANCE OF THE *IQ* CHANGES, FIRST TO SECOND TEST, FOR SEX GROUPS, KUHLMANN AND CATTELL, EXAMINERS *S* AND *H*, AND ORDER OF SCALE, DISREGARDING CHRONOLOGICAL AGE, TOTAL CASES, 32

Source	*df*	Sums of squares	*F*
Between Subjects:	(31)	(2297)	
Sex	1	39	
Scale Order	1	64	
Examiner	1	390	6.55*
Sex X Order	1	189	3.18
Sex X Examiner	1	182	3.06
Order X Examiner	1	3	
Order X Sex X Examiner	1	2	
Error (between)	24	1428	
Within Subjects:	(32)	(1378)	
Scales	1	18	
Scales X Sex	1	4	
Scales X Order	1	150	3.87
Scales X Examiner	1	31	
Scales X Sex X Order	1	121	3.12
Scales X Sex X Examiner	1	39	
Scales X Order X Examiners	1	63	
Scales X Order X Sex X Examiners	1	23	
Error (within)	24	929	
Total	63	3675	

*p. $< .05$.

b. Effects associated with age, sex, scale, scale order, and test period. The effects of these factors were studied by using the entire 80 cases, since examiner differences were counterbalanced for the subgroups. The data were analyzed again with a mixed design taking all five factors into consideration. The relevant means are given in Table 10, the rows representing independent subgroups, the columns, repeated measures on the same individuals. A summary of the analysis of variance is presented in Table 11. Only the simple (first-order) interactions are presented in this table, higher interactions having proved nonsignificant. The analysis indicated that the level of *IQ* for the same group depended on the scale used, the Kuhlmann mean being lower

TABLE 10

MEAN *IQs* FOR THE AGE, SEX, AND TEST-ORDER SUBGROUPS OF FIVE CASES EACH, FOR THE KUHLMANN AND CATTELL INTELLIGENCE SCALES, TEST AND RETEST

			Kuhlmann		Cattell	
Age groups	Sex	Order	Initial	Final	Initial	Final
I	M	KC	103.0	101.6	112.0	108.8
		CK	107.0	108.2	114.4	113.2
	F	KC	112.8	116.2	117.4	123.6
		CK	115.6	112.2	130.4	126.0
II	M	KC	105.8	102.2	124.2	127.6
		CK	103.0	98.6	111.2	111.6
	F	KC	105.2	112.8	128.6	134.4
		CK	107.2	110.6	125.2	125.8
III	M	KC	99.4	100.2	122.2	123.4
		CK	102.0	103.2	116.8	119.0
	F	KC	103.8	112.6	120.0	125.2
		CK	97.6	93.2	116.0	116.8
IV	M	KC	102.4	114.2	127.0	131.6
		CK	97.6	101.2	111.6	122.6
	F	KC	103.0	177.4	127.4	131.4
		CK	98.0	102.6	114.4	119.8

TABLE 11

SUMMARY OF ANALYSIS OF VARIANCE OF *IQs* FOR AGE, SEX, AND SCALE ORDER SUBGROUPS, FOR THE INITIAL AND FINAL TESTS ON KUHLMANN AND CATTELL, HIGHER ORDER INTERACTIONS OMITTED

Source	*df*	Sums of squares	*F*
Between Subjects:	(79)	(44501)	
Age	3	733	
Scale Order	1	1638	3.03
Sex	1	1911	3.54
Age X Order	3	2282	
Age X Sex	3	1756	
Sex X Order	1	10	
Error (between subjects)	67	36171	
Within Subjects:	(240)	(43696)	
Scales	1	20257	225.50***
Test Periods	1	573	6.38*
Scales X Periods	1	0	
Scales X Age	3	1412	5.24**
Scale X Order	1	97	
Scales X Sex	1	15	
Periods X Age	3	664	2.46
Periods X Order	1	214	2.38
Periods X Sex	1	72	
Error (within subjects)	227	20392	
Total	319	88197	

$^{***}p < .001$.
$^{**}p < .01$.
$^{*}p < .05$.

for every subgroup in Table 10. Similarly, it appears that there was a consistent gain from the first to the second test and that the amount of this gain was about the same for both scales. The significant test by age interaction represented the tendency for the difference between means on the two scales to increase with increasing age. Two other main and two additional interaction effects approached statistical significance at the 5 per cent level. These were sex, scale order, test period by age, and test period by scale order. Further discussion of these nonsignificant effects would appear unwarranted, despite the consistency of the trends which they represent.

2. *Factors Associated with Group Means for the Speech Indices*

Two 60 breath-unit records of speech were obtained for each child, one during each of the experimental sessions. It was found that the means of the phoneme type and phoneme frequency indices for these two sets of records differed but little, and not statistically significantly, for the whole group. Thus, it appeared reasonable to analyze scores based on a composite of these two records. The two sets of scores—phoneme type based on 120 breath-units and phoneme frequency based on the average of two 60 breath-unit samples—were analyzed to study the effects on them of the age and sex of subjects.

a. The phoneme type index. The means of the subgroups for this index are presented in Table 12. The data were ordered in a double-classification table of four age levels and the two sex groups, and were subjected to analysis of variance. The summary of the analysis is presented in Table 13. This table indicates that the age variable was significant beyond the 1 per cent level of confidence. There was an increase with age in the number of phoneme types produced. There was also a difference between the means for the sex groups which was significant beyond the 5 per cent level; girls tended to produce more different sounds at all age levels studied. These findings agree with those reported by Irwin and Chen (12).

TABLE 12
MEANS FOR AGE-SEX SUBGROUPS ON THE PHONEME TYPE AND PHONEME FREQUENCY INDICES

Age groups	Phoneme Type		Phoneme frequency	
	Boys	Girls	Boys	Girls
1	21.0	21.4	153.20	152.35
2	26.2	26.9	206.15	257.15
3	27.2	30.0	263.05	298.10
4	30.1	31.2	353.00	350.60

b. The phoneme frequency index. The data for this index were not so simply handled due to the heterogeneity of variance; the Bartlett test yielded a probability of less than .001 in the test of the homogeneity of the variances for the eight age-sex subgroups. Examination of the phoneme frequency means in Table 12, however, indicates an indubitable increase with increasing age. There is also an indication that the girls tended to have higher mean scores than did the boys, but this difference was small in comparison to the individual variability. An analysis of variance in a double-classification table (age and sex) showed the complex hypothesis tenable that the sex groups did not differ in means *and* variances. It thus appears that any further statistical analyses would be unprofitable.

TABLE 13
SUMMARY OF ANALYSIS OF VARIANCE OF PHONEME TYPE SCORES FOR THE AGE-SEX SUBGROUPS

Source	*df*	Sums of squares	*F*
Age	3	989.50	42.02***
Sex	1	31.25	3.98*
Age X Sex	3	17.25	
Individuals	72	565.00	
Total	79	1603.00	

***$p < .001$.
*$p < .05$.

F. DISCUSSION AND SUMMARY

Both intelligence scales used in the present study have been demonstrated to have fairly satisfactory degrees of consistency when used with young children under the conditions of the present study. The test-retest correlations ranged from .79 to .96 for groups of 20 children of different age levels. It seems important to emphasize, however, that the experimenters took definite steps to reduce disagreements in both scoring and administration and that under other conditions the reliabilities might be considerably lower. There appeared to be no evidence from this study that changing examiners from initial to final test materially affected the test-retest reliabilities.

Within the limits of the manual recording techniques used, the observer agreement or objectivity of speech sound recording appeared favorable. While point-by-point comparisons of the records of the two observers indicate that considerable disagreement occurs, comparisons of the overall frequencies of different sounds for a sample of speech sound units from a given child indicate respectable observer agreement. Improvement of recording methods seems desirable.

The test-retest correlations for the speech sound indices are not extremely high, ranging from .51 to .79 for groups of 20 children of homogeneous ages. Comparisons of split-half reliabilities and test-retest correlations indicate that at least part of the lack of consistency is associated with factors other than the recording method. That is, test-retest correlations tended to be lower than corrected split-half reliability estimates and therefore suggest that daily fluctuations in speech sound production may exist. One recommendation of the writers is that standard test stimuli be used to evoke the speech behavior in order to control some of the diurnal variability.

The intercorrelations of the two scales for the initial, retest, and average *IQ*s were quite high. The average correlation for the initial test was .78; for the retest, .80; and for the test-retest average, .85. These results indicate clearly that the level of performance of a child on one scale can be reliably predicted from his performance on the other scale. This does not imply, of course, that the *IQ*s obtained from the two scales will be approximately the same.

The relationships between the speech indices and the intelligence scales do not appear to be quite so simple as those between intelligence scales. The correlation between intelligence and speech sound production at the youngest age level appears to be negligible (or slightly negative). For the three older age levels, the relationship is positive and fairly high. Data from the present study do not provide evidence to explain the obtained relationship. For example, it is plausible to assume either that conditions favorable for the learning of speech sounds are also favorable for the learning of solutions to the types of problems found on intelligence scales, or that the same innate factors account for both. It is interesting to note that meaningful speech was extremely rare in the records of the youngest children. It is also possible that the speech indices are related to the facility with meaningful speech just as it is ordinarily assumed that the intelligence scales are positively correlated with such indices of speech development as spoken vocabulary, sentence structure, length of sentence, etc., among subjects over a year and a half in age.

Although test-retest correlations for the intelligence scales when one examiner gave both tests did not differ markedly from those obtained when different examiners administered the test and retest, significant differences were found between the mean *IQ*s obtained by the two examiners. Examiner *S* consistently obtained higher *IQ*s than did Examiner *H*. The mean difference was approximately three points for the Kuhlmann and approximately two points on the Cattell. These findings, especially considering the writers' attempts to establish examiner reliabilities, indicate clearly the danger of ignoring examiner differences in discussing a specific *IQ* for a given child.

Another error which these data point up is the frequently made assumption that a child who obtains a given *IQ* on one scale will make a similar score on the other scale. The Kuhlmann scale averaged approximately 16 *IQ* points lower on both test and retest than did the Cattell. Several instances were observed in which the Cattell provided an *IQ* over 30 points higher than that obtained from the Kuhlmann administered to the same child on the same day. One difference of 49 points was obtained. The mean difference in *IQ* scores provided by the two scales appeared to increase with age, at least within the age range studied, resulting in a significant scales by age interaction mean square.

An additional factor found to be significantly related to *IQ* level of groups was the frequently reported practice effect. That is, subjects gained on the average of about three points from the initial to final tests for both scales. This result appeared despite the examiners' avoidance of any teaching or coaching on the items during either session.

The data from the speech sound indices indicate clearly the results which have been reported by Irwin (8, 12), viz., that the phoneme type and phoneme frequency indices increase as a function of age. Irwin and Chen's (12) findings that girls tend to be superior to boys in the production of different sounds was substantiated. No sex differences were found for the phoneme frequency index.

References

1. Brodbeck, A. J., & Irwin, O. C. The speech behavior of infants without families. *Child Devel.,* 1946, **17**, 145-156.
2. Cattell, P. The Measurement of Intelligence of Infants and Young Children. New York: Psychological Corporation, 1940.
3. Chen, H. P., & Irwin, O. C. Infant speech: Vowel and consonant types. *J. Speech Disor.,* 1946, **11**, 27-29.
4. Fairbanks, G. Voice and Articulation Drillbook. New York: Harper, 1950.
5. Goodenough, F., & Maurer, K. The Mental Growth of Children from Two to Fourteen Years. Minneapolis: Univ. Minnesota Press, 1942.
6. Harms, I. E. A study of some variables affecting the reliability of intelligence test scores during late infancy. Unpublished Ph.D. Dissertation, State Univ. Iowa, 1951.
7. Irwin, O. C. The developmental status of speech sounds in ten feeble-minded children. *Child Devel.,* 1942, **13**, 29-39.
8. ———. The development of speech during infancy: The curve of phoneme frequency. *J. Exp. Psychol.,* 1947, **37**, 187-193.
9. ———. Infant speech: The effect of family occupational status and of age on the use of sound types. *J. Speech & Hear. Disor.,* 1948, **13**, 224-226.
10. Irwin, O. C., & Chen, H. P. Reliability study of speech sounds observed in the crying of newborn infants. *Child Devel.,* 1941, **12**, 351-368.
11. Irwin, O. C., & Chen, H. P. Infant speech sounds and intelligence. *J. Speech Disor.,* 1945, **10**, 293-296.

9. 3. 81
745

12. ———. Development of speech during infancy: Curve of phonemic types. *J. Exp. Psychol.,* 1946, **36**, 431-436.

13. ———. Infant speech: Vowel and consonant frequencies. *J. Speech Disor.,* 1946, **11**, 123-125.

14. Irwin, O. C., & Curry, T. Vowel elements in the crying vocalization of infants under ten days of age. *Child Devel.,* 1941, **12**, 99-109.

15. Kuhlmann, F. A Handbook of Mental Tests. Baltimore: Warwick & York, 1922.

16. Kuhlmann, F. Tests of Mental Development. Minneapolis: Educational Test Bureau, Educational Publishers, 1939.

17. Lindquist, E. F. Design and Analysis of Experiments in Psychology and Education. Chicago: Houghton Mifflin, 1953.

18. Maurer, K. M. Intellectual Status at Maturity as a Criterion for Selecting Items in Preschool Tests. Minneapolis: Univ. Minnesota Press, 1946.

19. Snedecor, G. Statistical Methods. Ames, Iowa: Iowa State College Press, 1946.

20. Spiker, C. C. An empirical investigation of the relationships between intelligence test scores and certain indices of speech sound development in young children. Unpublished Ph.D. Dissertation, State Univ. Iowa, 1951.

Iowa Child Welfare Research Station
State University of Iowa
Iowa City, Iowa

The Journal of Genetic Psychology, 1959, **94**, 23-27.

STABILITY OF ADOLESCENTS' JUDGMENTS OF AN AMBIGUOUS FIGURE AFTER KNOWLEDGE OF OTHERS' JUDGMENTS OF SAME FIGURE*

Department of Psychology, Grinnell College

Alice Van Krevelen

A. Problem

Adolescence is generally conceded to be a period during which the influence of the peer group on the individual's social behavior is intensified and the pressure to conform is especially strong. This experiment attempts to discover the extent to which girls aged 13, 14, and 15 will be influenced by the opinions of their peers and adults in a situation requiring *S* to make a judgment of an ambiguous figure both with and without knowledge of others' judgments of the same figure.

It should be of interest to investigate effects of group pressure in a situation of this type. In this particular study it was possible to present *S* with opinions both of her peers and of adults with whom she had been closely associated and who were in the position of leaders of the peer group. The question of whether adult opinion would carry any more weight in a problem situation such as the one used in this experiment than in purely social situations involving interaction in the peer group seemed to be one worth investigating.

B. Subjects and Procedure

Taking part in the investigation were 77 individuals in a girls' summer camp. There were 54 campers composed of approximately equal numbers of 13, 14, and 15 year olds and 23 counselors and counselors-in-training.

The stimulus was a large green figure on a white background. The figure was a straight line one and one fourth inches wide and nine inches long with prongs at either end, those at one end bending back above and below the line as an arrowhead and those at the other end continuing on the direction away from the line. The stimulus was placed on the mantle of a fireplace at one end of a large room where all *S*s were seated in a group facing it. To the left of the stimulus figure was the standard, a white line of the same width as the stimulus figure and about two thirds as long, on a green background.

---

*Received in the Editorial Office on October 21, 1957.

B[illegible] & Psyl. Rese
(. . E R I)
[illegible]te
Acc.

The experiment was introduced by *E*, a counselor, as part of an evening program and the directions were as follows:

> This is a contest to see how good you are at estimating the length of a line. There will be a prize for the camper who makes the closest estimate. Write your name on the slip of paper which was given you. Now look up here at the mantle. This line (pointing to the stimulus figure) is to be estimated in units of measure called gammas. This line over here at your left is 25 gammas long. Judging from that how long is the line on the right? By "line" I mean the horizontal part between the prongs. I think it only fair to tell you that these prongs distort the length of the line, those in front making it seem shorter than it actually is and those in back making it seem longer. Write on your slip of paper the number of gammas you estimate the line in the figure to be. Do not let anyone else see what you have written and do not tell anyone what you wrote until we announce the winner.

The slips of paper were then collected and the regular evening program started. At the end of an hour of activity (a scavenger hunt) the campers and counselors were once more assembled by *E* and the following announcement made:

> I am going to give you one more chance to guess the length of the line you looked at earlier now that you have had a little practice. We got quite varied results the first time. The counselors' guesses averaged about 24 gammas for the length of the line, the average for the counselors-in-training was 80 gammas, and the campers mostly guessed the line to be 41 gammas long. I will throw away your first paper and give you a chance to improve your original guess. Put your name on the new slip of paper and without looking at the figure again try to decide whether you over- or underestimated the length. If you think you were right the first time do not change your answer.

The mean figures reported to the group were arbitrarily chosen by *E* with the exception of the campers' mean which was the actual figure obtained by averaging the first judgments.

C. RESULTS

Subjects' first and second estimates were compared and directional changes determined. In general the change was considered to be in the direction of the reported mean to which the second estimate was closest. However there were some exceptions; for example if a judgment changed from 70 to 65 it was considered to be in the direction of the campers' mean even though the absolute magnitude of the second judgment was actually closer to the mean of 80 reported for the counselors-in-training. Likewise if a judgment was changed from 40 to 35 it was considered to be in the direction of the counselors' mean but if it changed from 30 to 35 it was considered to be in the direction of the campers' mean.

Of the camper group 56 per cent changed their judgments and 44 per cent did not. Of those who changed 24 per cent corrected estimates in the direction of the counselors' mean and 6 per cent in the direction of the mean reported for the counselors-in-training. Twenty six per cent changed to a figure closer to the reported mean of the other campers.

TABLE 1

PERFORMANCE OF *S*s AFTER KNOWLEDGE OF OTHERS' JUDGMENTS OF FIGURE

	Campers		Counselors-in-training		Counselors	
	Number	Per cent	Number	Per cent	Number	Per cent
No change in judgment	24	44	7	70	10	77
Change toward counselors' mean	13	24	1	10	3	23
Change toward counselors-in-training	3	6	2	20	0	0
Change toward campers' mean	14	26	0	0	0	0
Totals	54	100	10	100	13	100

Thirty per cent of the counselors-in-training changed their judgments and 70 per cent did not; 20 per cent of the changes were toward the campers' mean and 10 per cent closer to the mean reported for the counselors.

Of the counselors who took part in the experiment 23 per cent changed their judgments and 77 per cent did not. All of the changed estimates were closer to the counselors' mean than either of the other two reported means. None of the members of any group who did not change their judgments had made original estimates that coincided with any of the reported means.

D. DISCUSSION

The results indicate that more campers changed their judgments than did not and those who did change seemed to be influenced just about equally by the judgments of their peers and of their counselors. The counselor-in-training group actually represented a transition group being closer in age to the campers but closer in function to the counselor group. As a group they had the least influence on judgmental changes.

Although the adults in the experiment were merely incidental to the problem as stated it is interesting to note the trend toward an increasing number of individuals keeping their original judgments as we go from the camper group to the counselor group or from the younger to the older *S*s.

Unfortunately the numbers in the individual categories are too small to furnish a basis for any valid conclusion concerning this trend.

It should also be stated that of the campers not changing their judgments four made original estimates within one point of the reported campers' mean and seven original judgments were within one point of the counselors' mean. These facts throw some doubt onto our interpretation of the data to mean that 44 per cent of the group was not influenced by others' opinions. A judgment agreeing so closely with a reported mean allowed the *S* practically no way to indicate her reaction except by leaving her estimate unchanged. Does an unchanged judgment that close to a reported mean indicate that the individual is being influenced by the particular group or merely confident of her own judgment? In other words, if her judgment had not been close to the reported mean would she have changed it? Since *S*s were not asked this question directly the data afford no answer to it.

If we tentatively interpret the unchanged judgment to indicate agreement with the respective means then 33 per cent of the campers were influenced by other campers and 37 per cent by the counselors. This difference is still too small to permit us to conclude that either one of these groups had an edge over the other in influencing campers' opinions.

We can conclude that more than half the group were motivated to change their judgments but that the judgments of the peer group were weighted no more strongly than those of the counselor group. It would be interesting to repeat the same type of experiment on a younger aged group of children to determine whether the peer group is especially influential during the adolescent period even in a problem situation such as this.

Perhaps an explanation of the fact that a control-group type of experimental design was not used here is indicated. This procedure was not considered feasible since it seemed impossible to hold everything constant except the independent variable in the two groups. Whatever excuse was given to motivate the control group to make a second judgment would operate as another independent variable and serve to confuse the results. Since no *S* reported any difficulty in remembering her first judgment at the time she made her second the use of a control group which served no other purpose than establishing how well *S*s could remember their first judgments seemed unnecessary.

Since the figure was not shown a second time it was felt that any judgmental changes would be due to knowledge of others' estimates and not to practice effects of making a second independent judgment of the figure.

The fact that the experimental design may be responsible for a certain amount of ambiguity in the interpretation of results must also be considered in the light of the aims of the experiment. It was desired to present *S*s with a *choice* of reference groups on which to base their second judgments. Designs which would have led to more clearcut interpretation would most likely have entailed the presentation of just one reference group norm at a time. For reasons of motivation and "face validity" it would not have been possible to repeat the experiment three different times. For reasons of the size of the sample available it would not have been practical to divide *S*s into three different groups and present each group with just one reference norm. And even if this could have been done without introducing other uncontrolled variables or reducing the *N* too drastically results would be further contaminated, for in such a situation with just one reference given we might be testing *only* the degree of confidence *S*s had in their judgments. It is conceded that experimental nicety has been sacrificed in order to keep the procedure convincing to *S*s and prevent their suspecting that it was anything but part of an evening program and to better satisfy the aims of the investigation.

E. SUMMARY

When the *S*s in this experiment were given an opportunity to change their judgments of the length of a line after hearing others' estimates of the same line it was found that 44 per cent kept their original judgments, 24 per cent were influenced by the judgments of adults, and 26 per cent by the opinions of their peers.

A trend was also noted in that the older individuals were more likely to keep their original estimates, 77 per cent of the counselor group leaving their estimates unchanged, and all changes made by the adult group being in the direction of the reported mean of the other adults.

We can then conclude that campers who changed their estimates were influenced about equally by the judgments of their peers and those of the counselors. The counselor-in-training group had the least influence on camper judgments.

Department of Psychology
Grinnell College
Grinnell, Iowa

The Journal of Genetic Psychology, 1959, **94**, 29-34.

FACTORS RELATING TO OPINIONS OF SEVENTH GRADE CHILDREN REGARDING THE ACCEPTABILITY OF CERTAIN BEHAVIORS IN THE PEER GROUP*

Department of Psychology, University of Oregon

STANLEY M. GOERTZEN[1]

A. INTRODUCTION

In this study we sought to discover the opinions of seventh graders toward certain behaviors of their peers, to determine the degree of acceptability of these behaviors, and to study factors of sex, age, birth order, intelligence, and social status related to their opinions.

In reviewing the literature it was found that there has been a variety of studies generally oriented around understanding the relative acceptability or unacceptability of children and their behavior. The most extensive studies of judgments of symptoms are made by Wickman, and follow-up studies by Ellis and Miller. The format used in these studies was not considered feasible in a study of children's opinions, however. The Ohio Social Acceptance and Recognition Scales, with their descriptions of behavior and their more clearly defined social distance points, provided considerable material for format. The Bogardus Social Distance Scale was considered and the technique but not the form itself is included in this study.

Research workers have used a variety of techniques in gathering data by using established scales and trait lists (with or without revision) or by devising their own. The conclusion seems to be that the investigation determines the instrument. There is a preponderence of studies based on lists of behavior items determined by adults (teacher-authors) rather than upon the ideas and wording of children themselves. It is apparent that investigators are in close agreement on the value of learning more about children's behavior, acceptance or rejection of it, and related factors; although very few practical recommendations, suggestions, or implications are given.

B. PRELIMINARY INVESTIGATIONS

Before undertaking the collection of data from the large samples of subjects needed for the major part of the research, several pilot investigations were

---

*Received in the Editorial Office on October 24, 1957.

[1] Doctoral dissertation completed at the University of Oregon in 1955. The author is indebted to Dr. John Pierce-Jones and Dr. Edward A. Taylor for their guidance and constructive criticism of this research project.

necessary. The first was to develop a research tool, an opinion scale, with descriptions of unacceptable behaviors and social distance points. Initially, 220 seventh grade children in Alameda County were asked to write descriptions of people or of things people their age did that made them dislike them or not select them as friends. From these writings a list of 16 behaviors in children's wordings was obtained. There were some 30 interviews held to make sure the final 30 items selected from the writings had clarity and common meanings for these children. They also provided the social distance points. The pilot form of the list of 30 behaviors was administered to two seventh-grade classes and two adult groups.[2] With the suggestions from discussions with these groups and the material from observations by the teachers who administered the pilot scale, the final form evolved, including 32 behaviors and a five-point social distance scale.

In this pilot study it was found that both students and adults used the entire range of social distance points and that there was a distribution with a definite central grouping on each item. The readability was checked and a few items slightly reworded so that the readability index according to the Lorge and Dale-Schall formulae was near the high third grade level.

The selection of schools was a fairly representative one with a wide range of socio-economic grouping. Teachers were given explicit instructions for the administration of the Opinion Scale and they administered them to 1,773 boys and girls in Alameda County, California. The completed Opinion Scales were gathered and classified according to sex, age, birth order, intelligence, and socio-economic status of the subjects in order that relevant group comparisons could be made. There were no sub-groups with fewer than 159 children.

C. Description of Total Group and Sub-Groups

The total group was 1,773 seventh grade students in 58 classes from 16 public schools in rural, suburban, and urban Alameda County, California. The median age for the group was 12 years and 10 months and the range was from 11 years and 3 months to 15 years and 6 months. The median intelligence was 103. The scores ranged from 51 to 173.

There were 898 boys and 875 girls in the total sample. The girls were two months younger than the boys and four *IQ* points higher at the median than the boys.

Three age groups were used: those above 13 years and 7 months, the middle group of those 12 years and 6 months to 13 years and 6 months, and the youngest group of those 12 years and 5 months and below.

---

[2] A graduate research seminar in pupil personnel under the guidance of Dr. Clifford Froelich.

Four groups were formed in the category of birth order: first born in the family of more than one child, second born, third in birth order or later born, and the only child. The first and second born were representative of the total group in all respects. There was a tendency for the only born to be more intelligent, younger, and to come from a high socio-economic level. Conversely, the later born child was more often from lower socio-economic levels and tended to be lower in *IQ*.

The 1,450 students who had current California Test of Mental Maturity scores were divided into three groups: (*a*) bright, 110 *IQ* and above, (*b*) average, 90 to 109 *IQ*, and (*c*) low, below 90 *IQ*. The socio-economic status of these groups tended to rise with the intelligence levels.

Three socio-economic groups were employed. The classification was based on the father's occupation and his educational level. Although each of the three intelligence levels was represented in the socio-economic groups, there were proportionately more bright children in the upper socio-economic group. There was a tendency for the socio-economically high children to be younger and for the socio-economically low children to be older. Of the 1,423 children classified into socio-economic groups, 10 per cent were in the upper group, 72 per cent were in the middle group, and 18 per cent were in the lower group.

D. SCORING

The completed scales were classified into the sub-groups according to the factors described above and a count was made of the responses of the five social distance positions for each of the 32 behavior items. In order to determine ranks, the social distance response positions were weighted as indicated in Table 1.

TABLE 1

	Position	Weight
No. 1	Have as friend	1
No. 2	Okay	2
No. 3	Not a friend	3
No. 4	Don't care for	4
No. 5	Dislike or hate them	5

The weighted scores were multiplied by the frequency of each position. The sum was then divided by the *N* for that particular group to obtain the mean rating of each item. The most disliked behaviors, those with the highest mean ratings, were ranked first; the least disliked behaviors, those with the lowest mean ratings, were ranked last. This was done for each sub-group and for the total group so that Spearman rank order correlations could be determined. These values were obtained using various appropriate pairs of sub-groups.

E. Analysis of Results

The behavior items are listed in Table 2 according to their rankings by the total group. The rank order of the items was much the same in each of the sub-groups with the exception of a very few items. Items 1 through 11 are least acceptable and defined by the children as "don't care for" and "dislike or hate." The second group consists of items ranked from 12 through 25, defined as somewhat unacceptable but not actively or completely rejected. The third group is made up of items ranked 26 through 32, defined by the children as behaviors found in members of the peer group whom they will accept as friends or consider "okay" but "not my friends."

The items were classified according to certain behavior syndromes following similar classifications of both Wickman and Griffiths. The "aggressive" and "delinquency related" behaviors were rated more objectionable than the "nonconforming" and "withdrawing" behaviors by all of the student sub-groups. The mean rating and ranking showed them to be in the following order, listed from least acceptable to least rejected: "delinquent," "aggressive," "nonconforming," and "withdrawing." For reference, the items ranked 1, 3, 4, 5, 12, and 19 are included in the "delinquency" group. The "aggressive" items are ranked 1, 2, 14, 17, and 18. The "non-compliant" items are ranked 5, 9, 14, 15, 21, and 22. The "withdrawing" items are ranked 9, 23, 25, 26, and 27. The syndrome groups were ranked almost identically by each of the sub-groups.

In general, the rankings of the behaviors according to sub-groups were so similar that comparisons yield no Spearman rho of less than .92. These are all statistically significant at the .001 level according to Festinger and Katz. Thus, it can safely be said that on the items of the Opinion Scale, using the social distance technique of opinion study, the differences between the various sub-groups are slight or negligible.

F. Discussion

The striking similarity in opinions of groups which differ in socio-economic levels, sex, intelligence, age, and birth order might not be expected in view of the thinking of some investigators who have contended that these groups are different not only in behavior and attitudes, but also in details of their

TABLE 2
THE 32 BEHAVIOR ITEMS LISTED ACCORDING TO THEIR RANKINGS BY THE TOTAL GROUP

Rank	Item	Behavior item
1	4	Some 7th graders hit, push and pick fights. They try to beat up on people.
2	10	Some 7th graders push others around. They are the "I'm the boss" kind.
3	5	Some 7th graders take things from desks. They take things that aren't theirs.
4	8	Some 7th graders swear and use bad words.
5	24	Some 7th graders smoke.
6	12	Some 7th graders don't play fairly and won't take turns. They quit if they don't have their own way. They are poor losers.
7	22	Some 7th graders talk about others. They talk a lot about other people's business.
8	2	Some 7th graders act big and show off. They act stuck up.
9	16	Some 7th graders think they should be first and have the best.
10	20	Some 7th graders tell on other people.
11	26	Some 7th graders go around with people who get in trouble.
12	3	Some 7th graders don't tell the truth.
13	25	Some 7th graders act as if they never make mistakes.
14	9	Some 7th graders talk loud and butt in. They bother others when they are working.
15	31	Some 7th graders copy school work.
16	30	Some 7th grade boys act like girls. They like to play girls' games.
17	6	Some 7th graders find fault with what others do. They tell others they don't do things right.
18	13	Some 7th graders get mad easily and act mad at others. They have a bad temper.
19	11	Some 7th graders don't keep a promise very well. They say something and then don't do it.
20	27	Some 7th graders don't trust people. They question what others do and say.
21	18	Some 7th graders are sloppy and wear dirty clothes. They are not neat.
22	7	Some 7th graders fool around and don't tend to business.
23	19	Some 7th graders ask others to help them a lot. They don't do things on their own.
24	32	Some 7th grade girls act like boys. They like to play boys' games.
25	21	Some 7th graders act unhappy and sad. They don't enjoy things.
26	15	Some 7th graders don't pay attention. They seem to be day-dreaming.
27	1	Some 7th graders are shy. They don't try to make friends or play with other people.
28	14	Some 7th graders are poor in games and sports and can't dance well.
29	17	Some 7th graders are not smart. They can't do their school work very well.
30	29	Some 7th grade girls go out with boys. They may go to a show or something.
31	28	Some 7th grade boys go out with girls. They may take them to a show or something.
32	23	Some 7th graders always do what is right. They don't get into trouble.

prior socialization. The present study does not support the above contentions. This study was meant to be an extension of our knowledge, however, and was not directed explicitly to the testing of such theories. Nevertheless, it is interesting to observe the great similarity in the expressed opinions of all of the sub-groups and to hypothesize a similar system of values operating in the rejection of certain behaviors exhibited by members of their group. Further research would seem merited which would clarify the bases of the similarities demonstrated by this investigation.

REFERENCES

1. BOGARDUS, E. S. Social distance and its practical implications. *Social. & Soc. Res.*, 1938, **22**, 462-476.
2. DALE, E., & SCHALL, J. S. A formula for predicting readability. *Educ. Res. Bull.*, 1948, **27**, 11-20.
3. ELLIS, D. B., & MILLER, L. W. Teachers' attitudes and behavior problems. *J. Educ. Psychol.*, 1936, **27**, 501-511.
4. FESTINGER, L., & KATZ, D., *Eds.* Research Methods in the Behavioral Sciences. New York: Dryden, 1953.
5. GRIFFITHS, W. Behavioral Difficulties of Children as Perceived and Judged by Parents, Teachers, and Children Themselves. Minneapolis: Univ. Minnesota Press, 1952.
6. LORGE, I. Predicting readability. *Teach. Coll. Rec.*, 1944, **45**, 404-419.
7. OHIO STATE DEPARTMENT OF EDUCATION. Ohio Recognition Scale, Form G-4-B. Columbus, Ohio: Ohio State University, 1946.
8. ———. Ohio Social Acceptance Scale for Intermediate Grades. Columbus, Ohio: Ohio State University, 1946.
9. WICKMAN, E. K. Children's Behavior and Teachers' Attitudes. New York: Commonwealth Fund, 1948.

2550 Buchanan
San Francisco, California

The Journal of Genetic Psychology, 1959, **94**, 35-42.

SOLUTION OF A BI-MANUAL COORDINATION PROBLEM BY MONKEYS AND CHIMPANZEES*

Department of Psychology, University of Texas, and Yerkes Laboratories of Primate Biology

ARNOLD A. MCDOWELL AND HENRY W. NISSEN[1]

A. PROBLEM

This paper describes the performances of monkeys and of young and adult chimpanzees on the "elevator detour problem." This problem was presented to all animals in three degrees of difficulty and the conditions of testing were highly similar.

The elevator problem, in common with most detour problems which have been used for studying "insightful" learning, allows the subject to survey visually all relevant aspects of the situation. It differs from most previously reported detour problems in that it requires the sustained coördinated use of two hands.

B. SUBJECTS, APPARATUS, AND PROCEDURE

Twenty-four male rhesus monkeys 3 to 9 years old, two adult female chimpanzees (Dolly and Pitipie), and three juvenile (pre-adolescent) chimpanzees 5 to 7 years old (Easter, Debi, and Portia) were used as subjects in this experiment. The experimental histories of the monkeys were nearly identical. The experimental histories of the chimpanzees were roughly equivalent to those of the monkeys. Both monkeys and chimpanzees had had previous experience on the bentwire problem (1).

The monkeys were restrained in the holding cage of a *WGTA* during testing. The basic test apparatus, shown in Figure 1, consisted of a ladder obstruction and the elevator. The vertical ladder, with ⅛ inch steel rungs spaced 4-inches apart center to center, was seated in the test table 5-inches from the front of the restraining cage. The base of the elevator was fastened

---

*Received in the Editorial Office on October 25, 1957.

[1] The experiments reported in this paper constituted parts of more extensive studies of the effects of irradiation with X- or gamma-rays on the behavior of primates. The chimpanzees were tested in 1954-55 at the Yerkes Laboratories of Primate Biology as part of a program supported by the United States Atomic Energy Commission, Contract AT-(40-1)-1553. Work with the monkeys was conducted in 1956 at the Radiobiological Laboratory of The University of Texas and the United States Air Force under Contract AF 18(600)-165. All data reported here were obtained before irradiation.

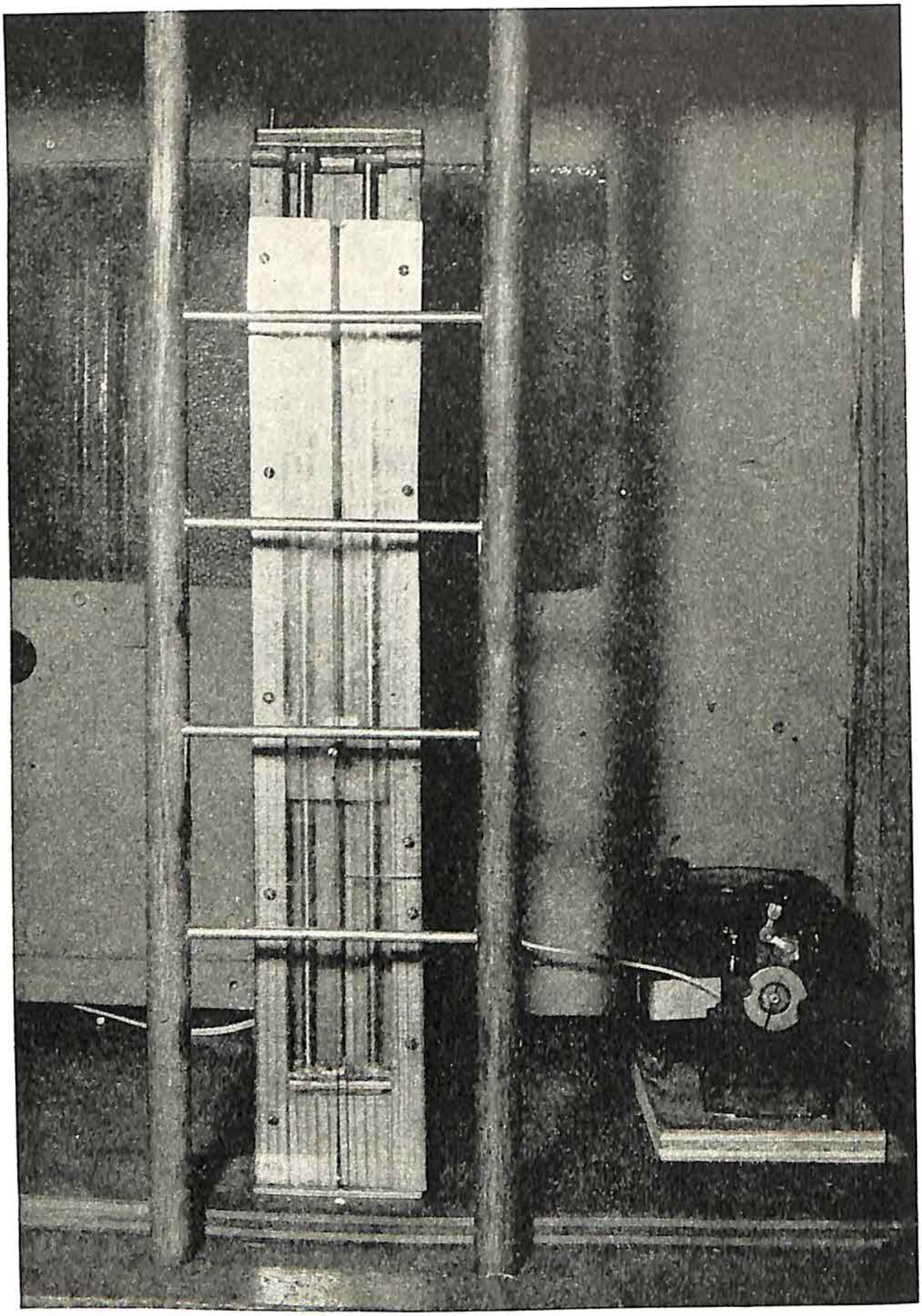

FIGURE 1
THE ELEVATOR PROBLEM AS VIEWED THROUGH THE LADDER OBSTRUCTION

to the test table floor behind the ladder, 7-inches from the front of the restraining cage.

To procure the food reward, the subject has to reach between successive ladder rungs to move a stylus vertically upward in a slot for distances of either 6, 12, or 18 inches. That is, the stylus is started 6, 12, or 18 inches from the top. As the stylus is raised, the lure can be seen through the Plexi glass front; when the stylus reaches the top of the slot, the lure becomes available to the subject. Since the stylus drops by gravity, the problem requires the coördinated use of two hands, one hand must hold the stylus while the other hand reaches over the next higher rung.

Prior to initiation of testing, each monkey was trained to the criterion of five successive procurements of the lure from each of the three vertical distances without the ladder obstruction in place. After meeting the adaptation criterion, each monkey was tested for 10 days with the ladder obstruction in place.

During each test day, each subject was tested for three trials on each of the three vertical distances. Order of presentation of distances was randomized. Dependent variables were approach latency from time of presentation of the problem, and performance time after initiating response. If a subject failed to initiate performance within 60 seconds after presentation, the trial was terminated and the subject was accorded a "failure without trying." If a subject failed to procure the lure within 60 seconds after initiating response, the trial was terminated and the subject was accorded a "failure with trying."

The chimpanzees were tested individually in their home cages, the 2-inch wire mesh of the cages serving a function analogous to that of the ladder obstruction with the monkeys. Under these conditions of testing, adaptation training without the obstruction in place could not be given. In all other respects the experimental procedure with the chimpanzees was identical to that with the monkeys.

C. RESULTS

1. *Test Reliability*

The number of failures to procure the lure on all trials of odd-numbered test-days by the 24 monkeys were enumerated and the subjects were ranked in order according to these scores. A similar ranking was based on the scores obtained by each subject on the even-numbered test-days and a rank order, Rho, correlation of 0.72 was obtained between the two sets of ranks. The reliability coefficient of the scores on the total test, is then, 0.83, as estimated by the Spearman-Brown Prophecy Formula.

2. *Species Difference*

Figure 2 shows the percentage of failures to procure the lure by both monkeys and chimpanzees on successive two-day test periods. Functions are presented for the 6-inch distance (Distance I), the 12-inch distance (Distance II), the 18-inch distance (Distance III), and all distances considered together. From left to right, each curve indicates the effects of increasing practice.

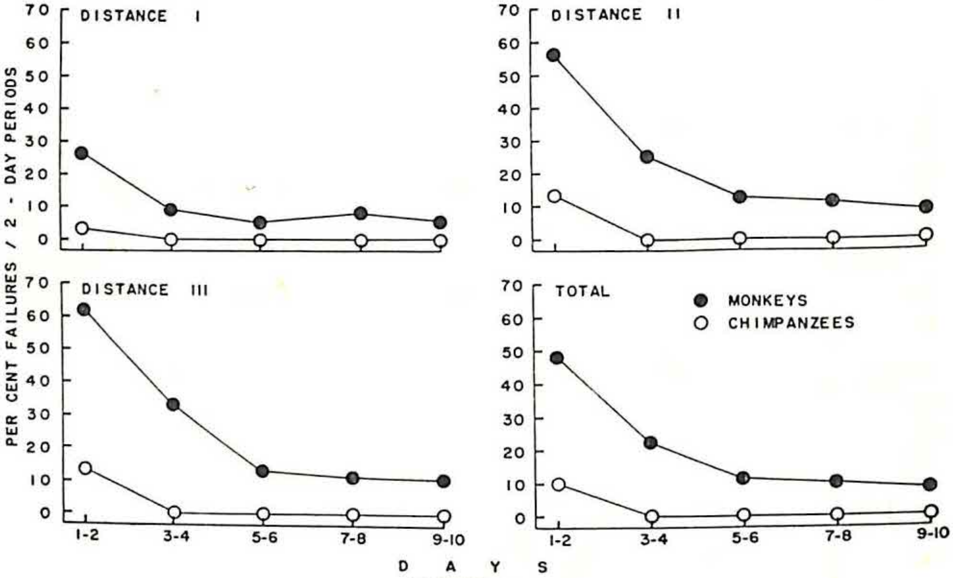

FIGURE 2

Percentage of Total Failures by Both Monkeys and Chimpanzees on Successive Two-Day Test Periods for Each Elevator Distance Separately for All Distances Considered Together

Differences between the performances of monkeys and chimpanzees are consistent and significant for all distances, the monkeys scoring more failures than the chimpanzees. As tested by White's rank test (2) for the significance of difference between groups, the differences between the 24 monkeys and the 5 chimpanzees are significant beyond the 5 per cent level of confidence for Distance I and beyond the 1 per cent confidence level for all other comparisons.

3. *Sources of Failure*

Only two of the chimpanzees recorded failures, these being "failures with trying." The monkeys, in contrast, scored "failures with trying" and "failures without trying" on all three distance conditions.

As tested by the sign test (6), the differences between total failures on each of the three distance conditions for the monkeys are significant beyond the 1 per cent confidence level, the number of failures increasing as the distance condition increases. Comparisons of the three distance conditions with respect to "failures to try" and "failures with trying," however, yield different results, as seen in Figures 3 and 4.

Figure 3 shows the percentage of failures to respond to the different distance

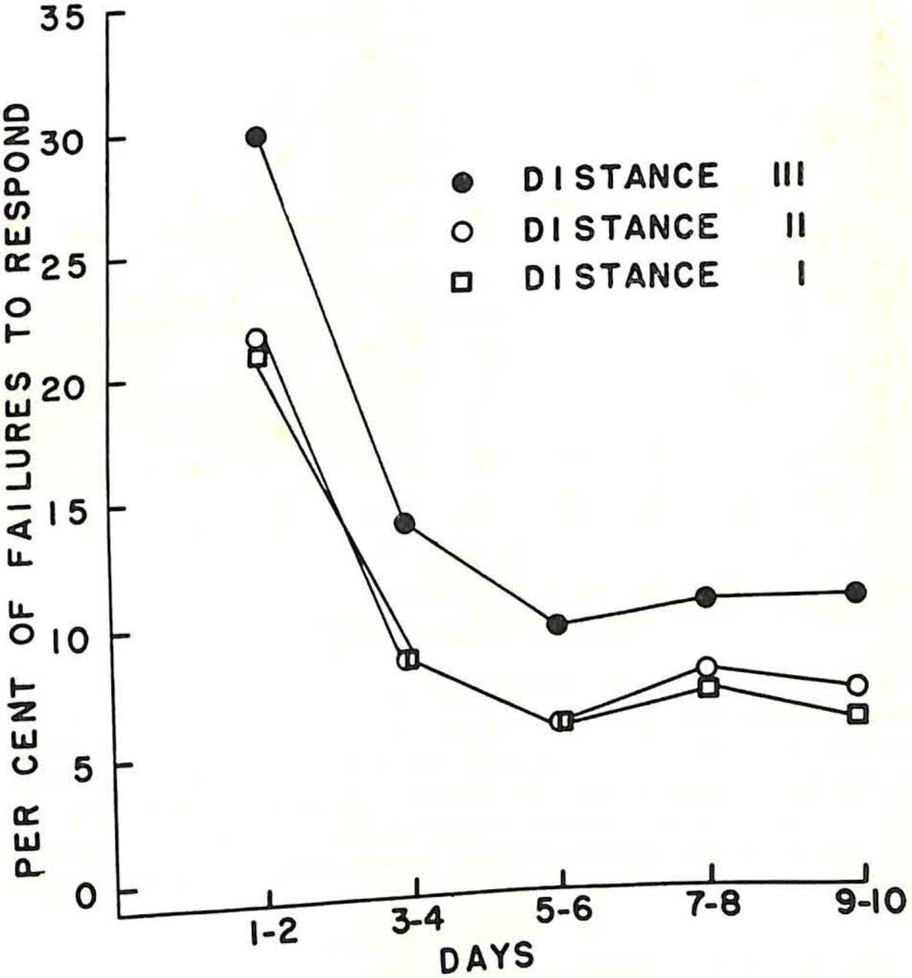

FIGURE 3

Percentage of Failures by Monkeys to Respond to the Different Distance Conditions During Successive Two-Day Periods—"Failures without Trying" conditions during successive two-day periods. As tested by the sign test, the frequency of failures to respond to Distance III is significantly greater than the frequency of failures to respond to Distance II ($p = 0.01$) and than the frequency of failures to respond to Distance I ($p = 0.02$). The frequencies of failures to respond to Distances I and II do not differ significantly.

Figure 4 shows the percentage of "failure with trying." Here, the frequencies of failures on Distances II and III do not differ significantly. The

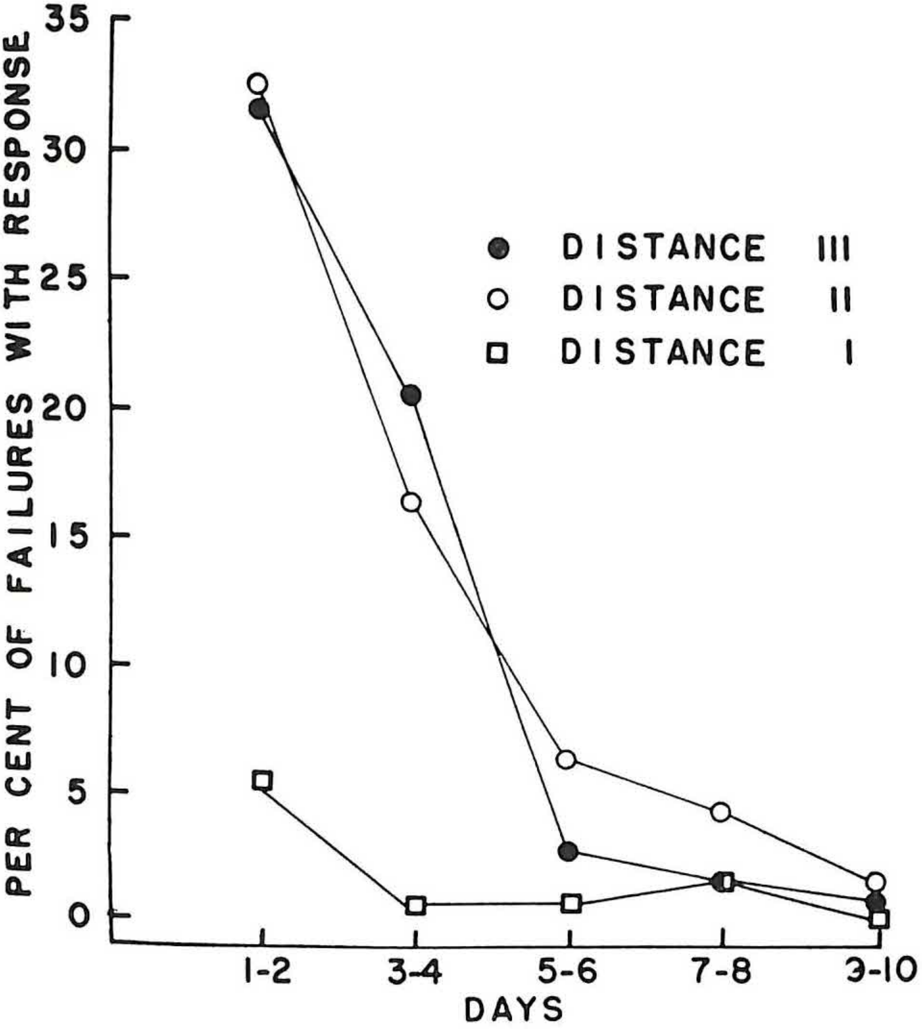

FIGURE 4
Percentage of Failures by Monkeys with Response to the Different Distance Conditions During Successive Two-Day Periods—"Failures with Trying"

frequency of failures on each of these two distances does differ significantly from the frequency of failures on Distance I ($p = 0.01$).

D. Discussion

The almost immediate mastery of the elevator detours by chimpanzees, in contrast to the gradual improvement in performance by monkeys agrees with differences found in patterned string tests by Finch (3) and Harlow and Settlage (4) and with differences found in bent-wire detour problems by Davis, McDowell, and Nissen (1).

The motor task was presumably more difficult for the chimpanzee than for the monkey, since the chimpanzee, with his larger hand, had to work through 2-inch wire mesh rather than through the 4-inch separations between ladder rungs used with the monkey. However, the fact that three of five chimpanzees had no failures, and the other two had failures only by reason of the performance time-limit, suggests that the species difference is a function more of perceptual factors than of superior motor skills.

The data suggest that, although the monkeys had some perception of the relative difficulty of the elevator distances employed, their perception of the difficulty of the intermediate distance was faulty. They appeared to perceive the intermediate distance as no more difficult than the shorter distance, whereas in terms of actual failures with response, it was as difficult as the longest distance condition. These results suggest an adaptation level phenomenon (5).

E. Summary

Twenty-four rhesus monkeys and five chimpanzees were tested on a new elevator detour problem. The performance of the chimpanzees was consistently superior to that of the monkeys. This species difference is interpreted as being primarily a function of general perceptual organization.

References

1. Davis, R. T., McDowell, A. A., & Nissen, H. W. Solution of bent-wire problems by monkeys and chimpanzees. *J. Comp. & Physiol. Psychol.*, in press.
2. Edwards, A. L. Statistical Methods for the Behavioral Sciences. New York: Rinehart, 1954.
3. Finch, G. The solution of patterned string problems by chimpanzees. *J. Comp. Psychol.*, 1941, **32**, 83-90.
4. Harlow, H. F., & Settlage, P. H. Comparative behavior of primates: VII. Capacity of monkeys to solve pattern string tests. *J. Comp. Psychol.*, 1934, **18**, 423-435.
5. Helson, H. Adaptation-level as a basis for a quantitative theory of frames of reference. *Psychol. Rev.*, 1948, **55**, 297-313.

6. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Radiobiological Laboratory
University of Texas
Austin, Texas

Yerkes Laboratories of Primate Biology
Orange Park, Florida

The Journal of Genetic Psychology, 1959, **94**, 43-66.

COGNITION OF BEING IN THE PEAK EXPERIENCES*[1]

Department of Psychology, Brandeis University

A. H. MASLOW[2]

A. INTRODUCTION

Self-actualizing people, those who have come to a high level of maturation, health, and self-fulfillment, have so much to teach us that sometimes they seem almost like a different breed of human beings. But, because it is so new, the study of the exploration of the highest reaches of human nature and of its ultimate possibilities and aspirations is a difficult and tortuous task. It has involved for me the continuous destruction of cherished axioms, the perpetual coping with seeming paradoxes, contradictions, and vaguenesses and the occasional collapse around my ears of long established, firmly believed in, and seemingly unassailable laws of psychology. Often these have turned out to be no laws at all but only rules for living in a state of mild and chronic psychopathology, and fearfulness, of stunting and crippling, and immaturity which we don't notice because most others have this same disease that we have.

Most frequently, as is typical in the history of scientific theorizing, this probing into the unknown first takes the form of a felt dissatisfaction, an uneasiness with what is missing long before any scientific solution becomes available. For instance, one of the first problems presented to me in my studies of self-actualizing people was the vague perception that their motivational life was in some important way different from all that I had learned. I first described it as being expressive rather than coping, but this wasn't *quite* right as a total statement. Then I pointed out that it was unmotivated rather than motivated, but this statement rests so heavily on which theory of motivation you accept, that it made as much trouble as help. In a more recent paper, I have contrasted growth-motivation with deficiency-need motivations which helps, I think, but isn't definitive enough yet, because it doesn't differentiate Becoming from Being. In this address, I shall propose a new tack (into a psychology of Being) which should include and generalize the three attempts already made to put into words somehow, the observed

---

*Received in the Editorial Office on October 29, 1957.

[1] Presidential Address, Division 8, A.P.A., Chicago, Ill., Sept. 1, 1956.

[2] I wish to thank James Klee, Frances Wilson Schwartz, and Harry Rand for helping me with this paper.

differences between the motivational and cognitive life of fully developed people and of most others.

This analysis of states of Being (temporary, non-striving, purposeless, self-validating, end-experiences and states) emerged first from a study of the love-relations and sexual experiences of self-actualizing people, and then of other people as well, and finally from dipping into the theological and philosophical literatures. It became necessary to differentiate two types of love. The one is love that comes from ordinary love-need, what Fenichel calls love as the need for narcissistic supplies, for gratification of a deficiency of love. It can therefore be conveniently called deficiency-love (D-love). It is typically and normally found in children and adolescents (of whatever age) in our culture.

But this creates a paradox. Self-actualizing people, by definition gratified in their basic needs, including the love need, should cease loving and wanting love, if the only determinant of love were the basic love-need. But the finding is that they are *more* loving people than the average, rather than *less* loving, especially from the point of view of being able to give love as well as to receive it. The attempt to understand this led to the formulation of another form or type of love, closely akin to what the theologians have called Agapean love, or Godly love, and which the psychoanalysts have named object-love and never described further. It is a love for the essence of or the Being of the other person, in the style that Scheler has described, quite apart from what he can give the lover, a love for the person in himself rather than for what we can get from him, detached, altruistic, admiring, unneeded, unselfish. It is love for another person because he is what he is, rather than because he is a need-gratifier.

In this state of love for the Being of the other person or object, I found a particular kind of cognition for which my knowledge of psychology had not prepared me but which I have since seen well described by certain writers on esthetics, religion, and philosophy. This I shall call Cognition of Being, or for short, B-Cognition. This is in contrast to cognition organized by the deficiency needs of the individual, which I shall call D-Cognition. The B-lover is able to perceive realities in the beloved to which others are blind, i.e., he is more acutely and penetratingly perceptive.

This paper is an attempt to generalize in a single description some of these basic cognitive happenings in the B-love experience, the parental experience, the mystic or oceanic, or nature experience, the aesthetic perception, the creative moment, the therapeutic or intellectual insight, the orgasmic experience, certain forms of athletic fulfillment. These and other moments of highest happiness and fulfillment I shall call the peak-experiences.

This is then a chapter in the Positive or Ortho-Psychology of the future in that it deals with fully functioning and healthy human beings, and not alone with normally sick ones. It is therefore not in contradiction to Psychology as a Psychopathology of the Average; it transcends it and can in theory incorporate all its findings in a more inclusive and comprehensive structure which includes both the sick and the healthy, both deficiency, Becoming and Being.

B. B-COGNITION IN PEAK EXPERIENCES

I shall present one by one now in a condensed summary, each of the characteristics of the cognition found in the generalized peak-experience, using the term cognition in an extremely broad sense.

1. *In B-Cognition the experience or the object tends to be seen as a whole, as a complete unit, detached from relations, from possible usefulness, from expediency, and from purpose.* It is seen as if it were all there was in the universe, as if it were all of Being, synonymous with the universe.

This contrasts with D-Cognition, which includes most human cognitive experiences. These experiences are partial and incomplete in ways that will be described below.

We are reminded here of the absolute idealism of the 19th century, in which all of the universe was conceived to be a unit. Since this unity could never be encompassed or perceived or cognized in any other fashion by a limited human being, all actual human cognitions were perceived as necessarily part of Being, and never conceivably the whole of it.

2. *When there is a B-Cognition, the percept is exclusively and fully attended to.* This may be called "total attention," or as Schachtel has called it, "focal attention." What I am trying to describe here is very akin to fascination or complete absorption. In such attention the figure becomes *all* figure and the ground, in effect, disappears, or at least is not importantly perceived. It is as if the figure were isolated for the time being from all else, as if the world were forgotten, as if the percept had become for the moment the whole of Being.

Since the whole of Being is being perceived, then all those laws obtain which would obtain if the whole of the cosmos could be encompassed at once. I shall discuss this further below.

This kind of perception is in sharp contrast to normal perception. Here the object is attended to simultaneously with attention to all else that is relevant. It is seen imbedded in its relationships with everything else in the world, and as *part* of the world. Normal figure ground relationships hold, i.e., both the ground and the figure are attended to, although in different ways. Furthermore in ordinary cognition, the object is seen not so much *per se* but as a member of a class, as an instance in a larger category. This kind of perception I have described as "rubricizing," and again would point out that this is not so much a full perception of all aspects of the objects or person being perceived, as it is a kind of taxonomy, a classifying, a ticketing off into one file cabinet or another.

To a far greater degree than we ordinarily realize, cognition involves also placing on a continuum. It involves a kind of automatic comparing or judging or evaluating. It implies higher than, less than, better than, taller than, etc.

B-Cognition is quite different. It may be called non-comparing cognition or non-evaluating or judging cognition. I mean this in the sense in which Dorothy Lee has described the way in which certain primitive peoples differ from us in their perceptions.

A person can be seen per se, in himself and by himself. He can be seen uniquely and idiosyncratically, as if he were the sole member of his class. This is what we mean by perception of the unique individual, and this is, of course, what all clinicians try to achieve. But it is a very difficult task, far more difficult than we are ordinarily willing to admit. However, it can happen, if only transiently, and it does happen characteristically in the peak experience. The healthy mother, perceiving her infant in love, approaches to this kind of perception of the uniqueness of the person. Her baby is not quite like anybody else in the world. It is marvelous, perfect, and fascinating (at least to the extent that she is able to detach herself from Gesell's norms and comparisons with neighbors' children).

Concrete perceiving of the whole of the object implies, also, that it is seen with "care." Contrariwise, "caring" for the object will produce the sustained attention, the repeated examination that is so necessary for perception of all aspects of the object. The caring minuteness with which a mother will gaze upon her infant again and again, or the lover at his beloved, or the connoisseur at his painting will surely produce a more complete perception than the usual casual rubricizing which passes illegitimately for perception. We may expect richness of detail and a many sided awareness of the object from this kind of absorbed, fascinated, fully attending cognition. This contrasts with the product of casual observation which gives only the bare bones of the experience, an object which is seen in only some of its aspects in a selective way and from a point of view.

3. While it is true that all human perception is a product of the human being and is his creation to an extent, still we can make some differentiation between the perception of *external objects as relevant to human concerns and as irrelevant to human concerns.* Self-actualizing people are more able to perceive the world as if it were independent not only of them but also of human beings in general. This also tends to be true of the average human being in his highest moments, i.e., in his peak experiences. He can then more readily look upon nature as if it were there in itself and for itself, and not simply as if it were a human playground put there for human purposes. He can more easily refrain from projecting human purposes upon it. In a word, he can see it in its own Being rather than as something to be used, or something to be afraid of, or to be reacted to in some other human way.

As one example, let us take the microscope which can reveal to us as we look at histological slides either a world of per se beauty or else a world of threat, danger, and pathology. A section of cancer seen through a microscope, if only we can forget that it is a cancer, can be seen as a beautiful and intricate and awe-inspiring organization. A mosquito is a wondrous object. Viruses under the electron miscroscope are fascinating objects (or, at least, they *can* be if we can only forget their human relevance).

B-Cognition, because it makes human-irrelevance more possible, enables us thereby to see more truly the nature of the object in itself.

4. One difference between B-Cognition and average cognition which is now emerging in my studies, but of which I am as yet uncertain, is that repeated *B-cognizing seems to make the perception richer.* The repeated experiencing of a face that we love or a painting that we admire makes us like it more, and permits us to see more and more of it in various senses. This we may call intra-object richness.

But this so far contrasts rather sharply with the more usual effects of repeated experiencing, i.e., boredom, familiarization effects, loss of attention and the like. I have found to my own satisfaction—although I cannot prove it to anyone else—that repeated exposures to what I consider a good painting will make the painting look more beautiful to people preselected as perceptive and sensitive, while repeated exposures to what I consider a bad painting will make it look less beautiful.

In this more usual kind of perception, where so frequently the initial perception is simply a classification into useful or not useful, dangerous or not dangerous, repeated looking makes it become more and more empty. The task of normal perception which is so frequently anxiety-based or motivation-determined, is fulfilled in the first viewing and thereafter the object or person, now that it has been catalogued, is simply no longer perceived. Poverty shows up in repeated experiencing; so, also, does richness. Furthermore not only does poverty of the percept show up in repeated looking, but also the poverty of the beholder.

I am becoming more convinced that one of the main mechanisms by which love produces a profounder perception of the intrinsic qualities of the love object than does non-love, is that love involves fascination with the love-object, and therefore repeated and intent and searching looking, seeing with "care." Lovers can see potentialities in each other that other people are blind to. Customarily we snicker and say "Love is blind," but we must now make room for the possibility that love may be under certain circumstances more perceptive than non-love. Of course this implies that it is possible in some sense to perceive potentialities which are not yet actual. I do not think that this is as difficult a research problem as it sounds. The Rorschach test in the hands of an expert is also a perception of potentialities which are not yet actualized. I think this is a testable hypothesis in principle.

5. American psychology, or more broadly, Western psychology, in what I consider to be an ethnocentric way, assumes that human needs, fears, and interests must always be determinants of perception. The "New Look" in perception is based upon the assumption that cognition must always be motivated. The further assumption is implied that cognition is a coping, instrumental mechanism, and that it must to some extent be egocentric. It assumes that the world can be seen *only* from the vantage point of the interests of the perceiver and that the experience must be organized around the ego as a centering and determining point.

I consider this point of view ethnocentric not only because it arises so clearly as an unconscious expression of the Western world outlook, but also because it involves a persistent and assiduous neglect of the writings of philosophers, theologians, and psychologists of the Eastern world, particularly of the Chinese, Japanese, and Hindus, not to mention Western writers like Goldstein and Angyal.

My findings indicate that in the normal perceptions of self-actualizing people and in the more occasional peak experiences of average people, *perception can be relatively ego-transcending, self-forgetful, egoless.* It can be unmotivated, impersonal, desireless, unselfish, not needed, detached. It can be object-centered rather than ego-centered. That is to say, that the perceptual experience can be organized around the object as a centering point rather than being based upon the ego. It is as if they were perceiving something that had independent reality of its own, and was not dependent upon the beholder. It is possible in the aesthetic experience or the love experience to become so absorbed and "poured into" the object that the self, in a very real sense, disappears. Some writers on aesthetics, mysticism, on motherhood, and on love have gone so far as to say that in the peak experience we may even speak of identification of the perceiver and the perceived, a fusion of what was two into a new and larger whole, a super-ordinate unit. This could remind us of some of the definitions of empathy and of identification, and, of course, at once opens up the possibilities of research in this direction.

6. *The peak experience is felt as a self-validating, self-justifying moment which carries its own intrinsic value with it.* That is to say it is an end in itself, what we may call an end-experience rather than a means-experience. It is felt to be so valuable an experience, so great a revelation, that even to attempt to justify it takes away from its dignity and worth. This is universally attested to by my subjects as they report their love experiences, their mystic experiences, their aesthetic experiences, their creative experiences, and their bursts of insight. Particularly with the moment of insight in the therapeutic situation does this become obvious. By virtue of the very fact that the person defends himself against the insight, it is therefore by definition painful to accept. Its breaking through into consciousness is customarily crushing to the person. And yet, in spite of this fact, it is universally reported to be worth while, desirable and wanted. Seeing is better than being blind, even when seeing hurts. It is a case in which the intrinsic self-justifying, self-validating worth of the experience makes the pain worthwhile. Not only do my subjects attest to this finding but so, also, do the numerous writers on aesthetics, religion, creativeness, and love. Uniformly they describe these experiences not only as valuable intrinsically, but as *so* valuable that they make life worth while by their occasional occurrence. The mystics have always affirmed this great value of the great mystic experience which may come only two or three times in a lifetime.

The contrast is very sharp with the ordinary experiences of life, especially in the West, and, most especially, for American psychologists. Behavior is so identified with means to ends that by many writers the words "behavior" and "instrumental behavior" are taken as synonymous. Everything is done for the sake of some further goal, *in order to* achieve something else. The apotheosis of this attitude is reached by John Dewey in his theory of value, in which he finds no ends at all but only means to ends. Even this statement is not quite accurate because it implies the existence of ends. Rather to be quite accurate he implies that means are means to other means, which in turn are means, and so on ad infinitum.

The peak experiences are for my subjects ultimate goals of living and the ultimate validations and justifications for it. That the psychologist should by-pass them or even be officially unaware of their existence, or what is even worse, in the objectivistic psychologies, deny a priori the possibility of their existence as objects for scientific study, is incomprehensible.

7. In *all the common peak experiences, or at least in those which I have studied, there is a very characteristic disorientation in time and space.* This goes so far that it would be more accurate to say that in these moments the person is outside of time and space subjectively. For instance, in the creative furor, the poet or artist becomes oblivious of his surroundings, and of the passage of time. It is impossible for him when he wakes up to judge how much time has passed. Frequently he has to shake his head as if emerging from a daze to rediscover where he is.

But more than this is the frequent report especially by lovers of the complete loss of extension in time. Not only does time pass in their ecstasies with a frightening rapidity so that a day may pass as if it were a minute but also a minute so intensely lived may feel like a day or a year. It is as if they had, in a way, some place in another world in which time simultaneously stood still and moved with great rapidity. For our ordinary categories, this is of course a paradox and a contradiction. And yet this is what is reported and it is therefore a fact that we must take account of. I see no reason why this kind of experiencing of time should not be amenable to experimental research. The judgment of the passing of time in peak-experience must be very inaccurate. So, also, for consciousness of surroundings. This, too, must be much less accurate than in normal living and, therefore, can be researched with.

8. I have been much impressed with the implications of my findings for a psychology of values. I find them very puzzling and yet so uniform that it is necessary not only to report them but also to try somehow to understand them. To start at the end first, *the peak experience is only good and desirable, and is never evil or painful or undesirable.* The experience is intrinsically valid and the experience is perfect, complete, and needs nothing else. It is sufficient to itself. It is felt as being intrinsically necessary and inevitable. It is just as good as it *should* be. It is reacted to with awe, wonder, amazement, humility, and even reverence, exaltation, and piety. The word sacred is occasionally used to describe the person's reaction to it.

The philosophical implications here are tremendous. If, for the sake of argument, we accept the thesis that in the peak experience the nature of reality *may* be seen more clearly and its essence penetrated more profoundly, then this is almost the same as saying what so many philosophers and theologians have affirmed, that the whole of Being is only neutral or good, and that evil or pain or threat is only a partial phenomenon, a product of not seeing the world whole and unified.

Another way of saying this is to compare it with one aspect of the concept of God which is very widespread in many religions. Those Gods who can contemplate and encompass the whole of Being and who therefore understand it, must see it as good, just, inevitable, and must see "evil" as a product of limited vision and understanding. If we could be God-like in this sense then we, too, out of universal understanding would never blame or condemn or be disappointed or shocked. Our only possible emotions would be pity, charity, kindliness, and perhaps sadness for the shortcomings of the other. But this is precisely the way in which self-actualizing people react to the world, and in which *all* of us react in our peak moments. I remind you that this is precisely the way in which all psychotherapists *try* to react to their patients. We must grant, of course, that this god-like, universally tolerant and accepting attitude is extremely difficult to attain, probably even impossible in a pure form, and yet we know that this is a relative matter. We can approximate it more closely or less closely and it would be foolish to deny the phenomenon simply because it comes rarely and impurely. Though we can never be gods in this sense, we can be more godlike or less godlike.

In any case the contrast with our ordinary cognitions and reactions is very sharp. Ordinarily we proceed under the aegis of means-values i.e., of usefulness, desirability, badness or goodness, of suitability for a purpose. We evaluate, judge, condemn, or approve. We react to the experience in personal terms and perceive the world in reference to ourselves and our ends, thereby making the world no more than means to our ends. This is the opposite of being detached from the world, which means in turn that we are not really perceiving *it,* but perceiving ourselves in it. We perceive then in a deficiency-motivated way and can therefore perceive only D-values. This is different from perceiving the whole world, or that portion of it which in the peak experience we take as surrogate for the world. Then and only then can we perceive *its* values rather than our own. These I call the values of Being, or for short, the B-values. These are the same as Robert Hartman's "intrinsic values."

These B-values are so far as I can make out at this point, (*a*) wholeness, integration, unity and interconnectedness (*b*) necessity, perfection, completeness and inevitability (*c*) aliveness, good functioning, spontaneity and process (*d*) richness, intricacy, and complexity (*e*) beauty, awe-fulness (*f*) goodness, rightness, desirability (*g*) uniqueness, idiosyncracy and expressiveness (*h*) effortlessness, ease of achievement, lack of strain or striving, and finally (*i*) occasionally, but not always an element of humor or playfulness.

Not only is this, then, a demonstration of fusion and unity of the old trinity of the true, the good, and the beautiful, but it is more than that. I have elsewhere reported my finding that truth, goodness, and beauty are in the normal person only fairly well correlated with each other, and in the neurotic even less so. It is only in the healthy and mature human being, in the self-actualizing, fully functioning person that they are so highly correlated that for all practical purposes they may be said to fuse into a unity. I would now add that this is also true for other people in their highest moments i.e., in their peak experiences of love, of sex, of creativity, of aesthetic perception, of religious or mystic experience, and of insight and understanding.

This finding, if it turns out to be correct, is in direct and flat contradiction to one of the basic axioms that guides all scientific thought, namely, that the more objective and impersonal perception becomes, the more detached it becomes from value. Fact and value have almost always (by intellectuals) been considered to be antonyms and mutually exclusive. But perhaps the opposite is true, for when we examine the most ego-detached, objective, motivationless, passive cognition, we find that it claims to perceive values directly, that values cannot be shorn away from reality and that the most profound perceptions of "facts" are tinged with wonder, admiration, awe and approval, i.e., with value.

9. Normal experience is imbedded in history and in culture as well as in the shifting and relative needs of man. It is organized in time and in space. That is to say it is part of larger wholes and therefore is relative to these larger wholes and frames of reference. Since it is felt to depend upon man for whatever reality it has, then if man were to disappear, *it,* also, would disappear. Its organizing frames of reference shift from the interests of the person to the demands of the situation, from the immediate in time to the past and the future and from the here to the there. In these senses experience and behavior are relative.

Peak experiences are from this point of view more absolute and less relative. Not only are they timeless and spaceless in the senses which I have indicated above, not only are they detached from the ground and perceived more in themselves, not only are they relatively unmotivated and detached from the interests of man, but they are also perceived and reacted to as if they were in themselves, "out there," as if they were perceptions of a reality independent of man and persisting beyond his life. It is certainly very difficult and also very dangerous scientifically to speak of relative and absolute and I am quite aware that I am walking into a semantic swamp. And yet I am compelled by the many introspective reports of my subjects to report this differentiation as a finding with which we psychologists will ultimately have to make our peace. These are the words that the subjects themselves use in trying to describe experiences which are essentially ineffable. *They* speak of "absolute," *they* speak of "relative," and it is my duty to report it.

Again and again we ourselves are tempted to this kind of vocabulary, for instance, in the realm of art. A Chinese vase may be perfect in itself, may be simultaneously 2000 years old and yet fresh in this moment, universal rather than Chinese. In these senses at least it is absolute, even though also simultaneously relative to time, to the culture of its origin, and to the aesthetic standards of the beholder. Is it not meaningful also that the mystic experience has been described in almost identical words by people in every religion, every era, and in every culture. No wonder Aldous Huxley has called it "The Perennial Philosophy." The great creators, let us say as anthologized by Brewster Ghiselin, have described their creative moments in almost identical terms, even though they were variously poets, chemists, sculptors, philosophers, and mathematicians.

The concept of absolute has made difficulty partly because it has almost always been permeated with a static taint. It is now clear from the experience of my subjects that this is not necessary or inevitable. Perception of an aesthetic object or a beloved face or a beautiful theory is a fluctuating, shifting process but this fluctuation of attention is strictly *within* the perception. Its richness can be infinite and the continued gaze can go from one aspect of the perfection to another, now concentrating on one side of it, now on another. A fine painting has many organizations, not just one, so that the aesthetic experience can be a continuous though fluctuating delight as it is seen, in itself, now in one way, now in another. Also it can be seen relatively in one moment, absolutely in the next. We needn't struggle over whether it is *either* relative *or* absolute. It can be both.

10. Ordinary cognition is a very active process. It is characteristically a kind of shaping and selection by the beholder. *He* chooses what to perceive and what not to perceive, he relates it to his needs and fears and interests, he gives it organization, arranging and re-arranging it. In a word, he works at it. Cognition is an energy consuming process. It involves alertness, vigilance, and tension and is, therefore, fatiguing.

B-Cognition is much more passive and receptive than active although, of course, it never can be completely so. The best descriptions that I have found of this "passive" kind of cognizing comes from Eastern philosophers especially from Lao-Tse and the Taoistic philosophers. Krishnamurti has an excellent phrase to describe my data. He calls it "choiceless awareness." We could also name it "desireless awareness." The Taoistic conception of "let be" also says what I am trying to say, namely, that perception may be undemanding rather than demanding, contemplative, rather than forceful. It can be humble before the experience, non-interfering, receiving rather than taking, it can let the percept be itself. I am reminded here, also, of Freud's description of "free floating attention." This, too, is passive rather than active, selfless rather than egocentric, dreamy rather than vigilant, patient rather than impatient. It is gazing rather than looking, surrendering and submitting to the experience.

I have also found useful a recent memorandum by John Shlien on the difference between passive listening and active, forceful listening. The good therapist must be able to listen in the receiving rather than the taking sense in order to be able to hear what is actually said rather than what he expects to hear or demands to hear. He must not impose himself but rather let the words flow in upon him. Only so can their own shape and pattern be assimilated. Otherwise one hears only one's own theories and expectations.

As a matter of fact we may say that it is this criterion, of being able to be receiving and passive, that marks off the good therapist from the poor one of whatever school. The good therapist is able to perceive each person in his own right, freshly and without the urge to taxonomize, to rubricize, to classify and pigeon hole. The poor therapist through a hundred years of clinical experience may find only repeated corroborations of the theories which he learned at the beginning of his career. It is in this sense that it has been pointed out that a therapist can repeat the same mistakes for 40 years and then call it "rich clinical experience."

An entirely different, though equally unfashionable, way of communicating the feeling of this characteristic of B-cognition, is to call it, with D. H. Lawrence and other Romantics, non-voluntary rather than volitional. Ordinary cognition is highly volitional and therefore demanding, prearranged, and preconceived. In the cognition of the peak-experience, the will does not interfere. It is held in abeyance. It receives and doesn't demand. We cannot command the peak-experience. It happens *to* us.

11. *The emotional reaction in the peak experience has a special flavor of wonder, of awe, of humility before the experience.* This sometimes has a touch

of fear (although pleasant fear) of being overwhelmed. My subjects report this in such phrases as "This is too much for me." "It is more than I can bear." "It is too wonderful." The experience may have a certain poignancy and piercing quality which may bring either tears or laughter and which may be paradoxically akin to pain, although this is a desirable pain which is often described as "sweet." This may go so far as to involve thoughts of death in a peculiar way. Not only my subjects but many writers on the various peak experiences have made the parallel with the experience of dying, that is an eager dying. A typical phrase might be: "This is too wonderful. I don't know how I can bear it. I could die now and it would be all right." Perhaps this is in part a hanging on to the experience and a reluctance to go down from this peak into the valley of ordinary existence. Perhaps it is in part, also, an aspect of the profound sense of humility, smallness, unworthiness before the enormity of the experience.

12. Another paradox with which we must deal, difficult though it is, is found in the conflicting reports of perception of the world. *In some reports, particularly of the mystic experience or the religious experience or philosophical experience, the whole of the world is seen as a unity, as a single rich entity. In other of the peak experiences, most particularly the love experience and the aesthetic experience, one small part of the world is perceived as if it were for the moment all of the world.* In both cases the perception is of unity. Probably the fact that the B-cognition of a painting or a person or a theory retains all the attributes of the whole of Being, i.e., the B-values, derives from this fact of perceiving it as if it were all that existed at the moment.

13. In another paper I have tried to demonstrate the substantial difference between the cognition that abstracts and categorizes (rubricizing) and the fresh cognition of the concrete, the raw, and the particular. This is the sense in which I shall use the terms abstract and concrete. They are not very different from Goldstein's terms. There I pointed out also that most of our cognitions (attendings, perceivings, rememberings, thinkings, and learnings) were abstract rather than concrete. That is, we mostly categorize, schematize, classify, and abstract in our cognitive life. We do not so much cognize the nature of the world as it actually is, so much as we do the organization of our own inner world outlook. Most of experience is filtered through our system of categories and rubrics, as Schachtel has also pointed out in his classical paper on "Childhood Amnesia and the Problem of Memory." I was led to this differentiation by my studies of self-actualizing people, *finding in them simultaneously the ability to abstract without giving*

up concreteness and the ability to be concrete without giving up abstractness. This adds a little to Goldstein's description because I found not only a reduction to the concrete but also what we might call a reduction to the abstract, i.e., a loss of ability to cognize the concrete. Since then I have found this same exceptional ability to perceive the concrete in good artists as well, even though not self-actualizing. More recently I find this same ability in ordinary people in their peak moments. They are then more able to grasp the percept in its own concrete idiosyncratic nature.

Since this kind of idiographic perceiving has customarily been described as the core of aesthetic perceiving, as for instance by Northrop, they have almost been made synonymous. For most philosophers and artists, to perceive a person concretely, in his intrinsic uniqueness is to perceive him aesthetically. I prefer the broader usage and think that I have already demonstrated that this kind of perception of the unique nature of the object is characteristic of *all* peak experiences, not only the aesthetic one.

I find it useful to understand the concrete perceiving which takes places in B-Cognition as a perception of all aspects and attributes of the object simultaneously or in quick succession. Abstracting is in essence a selection out of certain aspects only of the object, those which are of use to us, those which threaten us, those with which we are familiar, or those which fit our language categories. Both Whitehead and Bergson have made this sufficiently clear, as have many other philosophers since. Vivanti has phrased it well when he pointed out that abstractions to the extent that they are useful, are also false. In a word, to perceive an object abstractly means *not* to perceive some aspects of it. It clearly implies selection of some attributes, rejection of other attributes, creation or distortion of still others. We make of it what we wish. We create it. We manufacture it. Furthermore, extremely important is the strong tendency in abstracting to relate the aspects of the object to our linguistic system. This makes special troubles because language is a secondary rather than a primary process in the Freudian sense, because it deals with external reality rather than psychic reality, with the conscious rather than the unconscious. It is true that this lack can be corrected to some extent by poetic language but in the last analysis much of experience is ineffable and can be put into no language at all.

Let us take for example the perception of a painting or of a person. In order to perceive them fully we must fight our tendency to classify, to compare, to evaluate, to need, to use. The moment that we say this man is, e.g., a foreigner, in that moment we have classified him, performed an abstracting act and, to some extent, cut ourselves off from the possibility of seeing him as a unique and whole human being, different from any other one in the whole world. In the moment that we approach the painting on the wall to read the name of the artist, we have cut ourselves off from the possibility of seeing it with complete freshness in its own uniqueness. To a certain extent then what we call *knowing,* i.e., the placing of an experience in a system of concepts or words or relations, cuts off the possibility of full cognizing. Herbert Read has pointed out that the child has the "innocent eye," the ability to see something as if he were seeing it for the first time (frequently he *is* seeing it for the first time). He can then stare at it in wonder, examining all aspects of it, taking in all its attributes, since for the child in this situation, no attribute of a strange object is any more important than any other attribute. He does not organize it; he simply stares at it. In the similar situation for the adult, to the extent that we can prevent ourselves from only abstracting, naming, placing, comparing, relating, to that extent will we be able to see more and more aspects of the many sidedness of the person or of the painting. Particularly I must underline the ability to perceive the ineffable, that which cannot be put into words. Trying to force it into words changes it, and makes it something other than it is, something else *like* it, something similar, and yet something different than *it* itself.

It is this ability to perceive the whole and to rise above parts which characterizes cognition in the various peak experiences. Since only thus can one know a person in the fullest sense of the words, it is not surprising that self-actualizing people are so much more astute in their perception of people, in their penetration to the core or essence of another person. This is also why I feel convinced that the ideal therapist, who presumably should be able as a professional necessity, to understand another person in his uniqueness and in his wholeness, without presupposition, ought to be at least a fairly healthy human being. I maintain this even though willing to grant unexplained individual differences in this kind of perceptiveness, and that also therapeutic experience can itself be a kind of training in the cognition of the Being of another human being. This also explains why I feel that a training in aesthetic perceiving and creating could be a very desirable aspect of clinical training.

14. *At the higher levels of human maturation, many dichotomies, polarities, and conflicts are fused and resolved.* Self-actualizing people are simultaneously selfish and unselfish, Dionysian and Appolonian, individual and social, rational and irrational, fused with others and detached from others, and so on. What I had thought to be straight line continua, whose extremes

were polar to each other and as far apart as possible, turned out to be rather like circles or spirals, in which the polar extremes came together into a fused unity. So also do I find this as a strong tendency in the full cognition of the object. The more we understand the whole of Being, the more we can tolerate the simultaneous existence and perception of inconsistencies, of oppositions, and of flat contradictions. These seem to be products of partial cognition, and fade away with cognition of the whole. The neurotic person, seen from a godlike vantage point, can then be seen as a wonderful, intricate even beautiful unity of process. What we normally see as conflict and contradiction and dissociation can then be perceived as inevitable, necessary, even fated. That is to say if he can be fully understood, then everything falls into its necessary place and he can be aesthetically perceived and appreciated. All his conflicts and splits turn out to have a kind of sense or wisdom. Even the concepts of sickness and of health may fuse and blur when we see the symptom as a pressure toward health, or see the neurosis as the healthiest possible solution at the moment to the problems of the individual.

15. *The person at the peak is godlike not only in senses that I have touched upon already but in certain other ways as well, particularly in the complete and loving and uncondemning acceptance of the world and of the person,* however bad he may look at more normal moments. The theologians have long struggled with the terrible task of reconciling sin and evil and pain in the world with the concept of an all-powerful, all-loving, all-knowing God. A subsidiary difficulty has been presented by the task of reconciling the necessity of rewards and punishments for good and evil with this concept of all-loving, all-forgiving God. He must somehow both punish and not punish, both forgive and condemn.

I think we can learn something about the resolution of this dilemma from the study of self-actualizing people and from the comparison of the two broadly different types of perception discussed so far, i.e., B-perception and D-perception. B-perception is a momentary thing ordinarily. It is a peak, a high spot, an occasional achievement. It looks as if all human beings perceive most of the time in a deficiency way. That is, they compare, they judge, they approve, they relate, they use. This means that it is possible for us to perceive another human being alternately in two different ways, sometimes in his Being, as if he were the whole of the universe for the time being. Much more often, however, we perceive him as a part of the universe and related to the rest of it in many complex ways. When we B-perceive him, *then* we can be all-loving, all-forgiving, all-accepting, all-admiring, all-understanding. But these are precisely the attributes assigned to most conceptions of

God. In such moments we can then be godlike in these attributes. For instance, in the therapeutic situation we can relate ourselves in this loving, understanding, accepting, forgiving way to all sorts of people whom we normally fear and condemn and even hate—murderers, pederasts, rapists, exploiters, cowards.

It is extremely interesting to me that all people behave at times as if they wanted to be B-cognized. They resent being classified, categorized, rubricized. Ticketing off a person as a waiter or a policeman or a dame instead of as an individual often offends. We all want to be recognized and accepted for what we are in our fulness, richness, and complexity. If such an acceptor cannot be found among human beings, then the very strong tendency appears to project and create a godlike figure.

Another kind of answer to the "problem of evil" is suggested by the way in which our subjects "accept reality" as being-in-itself, in its own right. It is neither *for* man nor is it *against* him. It just is impersonally what it is. An earthquake which kills poses a problem of reconciliation only for the man who needs a personal God who is simultaneously all-loving and omnipotent and who created the world. For the man who can perceive and accept it impersonally and uncreated it presents no ethical or axiological problem, since it wasn't done "on purpose," to annoy him. He shrugs his shoulders and if evil is defined anthropocentrically, he simply accepts evil as he does the seasons and the storms. Of course it is much harder to achieve this attitude with human actions which are hurtful to him, but it is occasionally possible, and the more matured the man is, the more possible it is.

16. *Perception in the peak moment tends very strongly to be idiographic.* The percept, whether a person or the world or a tree or work of art, tends to be seen as a unique instance, and as the only member of its class. This is in contrast to our normal nomothetic way of handling the world which rests essentially on generalization and on an Aristotelian division of the world into classes of various sorts, of which the object is an example. The whole concept of classification rests upon general classes. If there were no classes the concepts of resemblance, of equality, of similarity and of difference would become totally useless. One cannot compare two objects which have nothing in common. Furthermore for two objects to have something in common means necessarily abstraction, e.g., such qualities as redness, roundness, heaviness, etc. But if we perceive a person without abstracting, if we insist upon perceiving all his attributes simultaneously and as necessary to each other, then we no longer can classify. Every whole person from this point of view or every painting or every bird or flower becomes the sole member of a class and must therefore be perceived idiographically.

17. *One aspect of the peak-experience is a complete, though momentary, loss of fear, anxiety, inhibition, defense and control, a giving up of renunciation, delay, and restraint.* The fear of disintegration and dissolution, the fear of being overwhelmed by the "instincts," the fear of death and of insanity, the fear of giving in to unbridled pleasure and emotion, all tend to disappear or go into abeyance for the time being.

It may be thought of as pure gratification, pure expression, pure elation. But since it is "in the world," it represents a kind of fusion of the Freudian "pleasure principle" and "reality principle." It is therefore still another instance of the resolution of normally dichotomous concepts at higher levels of psychological functioning.

We may therefore expect to find a certain "permeability" in people who have such experiences commonly, a closeness and openness to the unconscious, and a relative lack of fear of it.

C. OTHER CHANGES IN THE PERSON

In addition to all these changes in the cognition of the world (including attitudes toward it), all sorts of other changes occur in the person in the peak experience and afterward.

For one thing, not only the world but also he himself becomes more a unity, more integrated, and self-consistent. This is another way of saying that he becomes more completely himself, idiosyncratic, unique. And since he is so, he can be more easily expressive and spontaneous without effort. All his powers then come together in their most efficient integration and coördination, organized and coördinated much more perfectly than usual. Everything then can be done with unusual ease and lack of effort. Inhibition, doubt, control, self-criticism, diminish toward a zero point and he becomes the spontaneous, coördinated, efficient organism, functioning like an animal without conflict or split, without hesitation or doubt, in a great flow of power that is so peculiarly effortless, that it may become like play, masterful, virtuoso-like. In such a moment, his powers are at their height and he may be startled (afterwards) by his unsuspected skill, confidence, creativeness, perceptiveness, and virtuosity of performance. It is all so easy that it can be enjoyed and laughed with. Things can be dared that would be impossible at other times.

To put it simply, he becomes more whole and unified, more unique and idiosyncratic, more alive and spontaneous, more perfectly expressive and uninhibited, more effortless and powerful, more daring and courageous (leaving fears and doubts behind), more ego-transcending and self-forgetful.

But these are almost the same as the list of B-values already described above. Which is to say that as the essential Being of the world is felt by the perceiver to be cognized, so also does he concurrently come closer to his own Being, or to being himself, or to self-actualization, or to perfection in his own kind, etc. As the experience becomes more unified, so does he also become more unified, as it becomes richer, so does he, as it becomes more itself, so does he as well, and so on. He and the world become more like each other as they both move toward perfection. Perhaps this is what is meant by the well-known fusion of lovers, the becoming one with the world in the mystic experience, the feeling of absorbed fusion and unity with the work of art in the aesthetic experience, the feeling so often reported by our great creators that their words take hold of them and practically write themselves as if they were being dictated, the great philosophical insights in which one becomes *part* of the unity one experiences and merged into it. This is what Angyal is talking about in part, when he speaks of the trend to homonomy. Also relevant here is my conclusion that just those qualities which describe a good painting (Wilson's criteria) also describe the good human being, i.e., the B-values of wholeness, uniqueness, aliveness, richness, etc.

May I now attempt briefly to put all of this in another frame of reference which is more familiar to you, the psychoanalytic. Secondary processes deal with the real world outside the unconscious. Logic, science, common sense, good adjustment, enculturation, responsibility, planning, rationalism are all secondary process techniques. The primary processes were first discovered in neurotics and psychotics and then in children, and only recently in healthy people. The rules by which the unconscious works can be seen most clearly in dreams. Wishes and fears are the primary movers, the Freudian mechanisms, the primary techniques. The well adjusted, responsible, common sense man who gets along well in the real world must usually do this in part by turning his back on his unconscious and denying and repressing it.

For me, this realization came most keenly when I had to face the fact about 15 years ago that my self-actualizing subjects, picked because they were very mature, were at the same time, also childish. I called it "healthy childishness," a "second naivete." It has since been recognized by Kris and the ego-psychologists as "regression in the service of the ego," not only found in healthy people, but finally conceded to be a *sine qua non* of psychological health. Balint has recognized love to be a regression (i.e., the person who can't regress can't love). And, finally, the analysts agree that inspiration and great (primary) creativeness comes partly out of the unconscious, i.e., is a healthy regression, a temporary turning away from the real world.

Now what I have been describing here may be seen as a fusion of ego, id, super-ego, and ego-ideal, of conscious and unconscious, of primary and secondary processes, a synthesizing of pleasure principle with reality principle, a regression without fear in the service of the greatest maturity, a true integration of the person at *all* levels.

D. REDEFINITION OF SELF-ACTUALIZATION

In other words, any person in any of the peak experiences, takes on temporarily many of the characteristics which I found in self-actualizing individuals. That is, for the time they become self-actualizers. We may think of it as a passing characterological change if we wish, and not just as an emotional-cognitive-expressive state. Not only are these his happiest and most thrilling moments, but they are also moments of greatest maturity, individuation, fulfillment—in a word his healthiest moments.

This makes it possible for us to redefine self-actualization in such a way as to purge it of all its static and typological shortcomings, and to make it less a kind of all-or-none pantheon into which some rare people enter at the age of 60. We may define it as an episode, or a spurt in which the powers of the organism come together in a particularly efficient and intensely enjoyable way, and in which he is more integrated and less split, more open for experience, more idiosyncratic, more perfectly expressive or spontaneous, or fully functioning, more creative, more humorous, more ego-transcending, more independent of his lower needs, etc. He becomes in these episodes more truly himself, more perfectly actualizing his potentialities, closer to the core of his Being.

Such states or episodes can, in theory, come at any time in life to any person. What seems to distinguish those individuals I have called self-actualizing people, is that in them these episodes seem to come far more frequently, and intensely and perfectly than in average people. This makes self-actualization a matter of degree and of frequency rather than an all-or-none affair, and thereby makes it more amenable to available research procedures. We need no longer be limited to searching for those rare subjects who may be said to be fulfilling themselves most of the time. In theory at least we may also search *any* life history for episodes of self-actualization, especially those of artists, intellectuals and other especially creative people, of profoundly religious people, and of people experiencing great insights in psychotherapy, or in other important growth experiences.

E. The Question of External Validity

So far I have described a subjective experience in a phenomenological fashion. Its relationship to the external world is another matter altogether. Just because the perceiver *believes* that he perceives more truly and more wholly, is no proof that he actually does so. The criteria for judging the validity of this belief ordinarily lie in the objects or persons perceived or in the products created. They are therefore, in principle, simple problems for correlational research.

But in what sense can art be said to be knowledge? The aesthetic perception certainly has its intrinsic self-validation. It is felt as a valuable and wonderful experience. But so also are some illusions and hallucinations. And furthermore you may be aroused to an aesthetic experience by a painting which leaves me untouched. If we are to go at all beyond the private, the problem of external criteria of validity remains, just as it does with all other perceptions.

The same can be said for loving perception, for the mystic experience, for the creative moment, and for the flash of insight.

The lover perceives in the beloved what no one else can, and again, there is no question about the intrinsic value of his inner experience and of the many good consequences for him, for his beloved, and for the world. If we take as an example, the mother loving her baby, the case is even more obvious. Not only does love perceive potentialities but it also actualizes them. The absence of love certainly stifles potentialities and even kills them. Personal growth demands courage, self confidence, even daring, and non-love from the parent or the mate produces the opposite, self doubt, anxiety, feelings of worthlessness and expectations of ridicule, all inhibitors of growth and of self-actualization.

All personological and psychotherapeutic experience is testimonial to this fact that love actualizes and non-love stultifies, whether deserved or not.

The complex and circular question then arises here "To what extent is this phenomenon a self-fulfilling prophecy" as Merton has called it. A husband's conviction that his wife is beautiful, or a wife's firm belief that her husband is courageous, to some extent *creates* the beauty or the courage. This is not so much a perception of something that already exists as a bringing into existence by belief. Shall we perhaps consider this an example of perception of a potentiality, since *every* person has the possibility of being beautiful and courageous? If so, then this is different from perceiving the real possibility that someone may become a great violinist, which is *not* a universal possibility.

And yet, even beyond all this complexity, the lurking doubts remain to those who hope ultimately to drag all these problems into the domain of public science. Frequently enough, love for another brings illusions, the perceptions of qualities and potentialities that don't exist, that are not therefore truly perceived but created in the mind of the beholder and which then rest on a system of needs, repressions, denials, projections, and rationalizations. If love can be more perceptive than non-love, it can also be blinder. And the research problem remains to nag us, when is which? How can we select out those instances in which perception of the real world is more acute? I have already reported my observations at the personological level, that one answer to this question lies in the variable of the psychological health of the perceiver, in or out of the love relationship. The greater the health, the more acute and penetrating the perception of the world, all other things being equal. Since this conclusion was the product of uncontrolled observation, it must be presented only as a hypothesis awaiting controlled research.

In general, similar problems confront us in aesthetic and intellectual bursts of creativeness, and also in the insight experiences. In both instances, the external validation of the experience is not perfectly correlated with phenomenological self-validation. It is possible for the great insight to be mistaken, the great love to disappear. The poem that creates itself in a peak experience may have to be thrown away later as unsatisfactory. Creation of a product that will stand up feels subjectively the same as the creation of a product that folds up later under cold, objective critical scrutiny. The habitually creative person knows this well, expecting half of his great moments of insight not to work out. All peak experiences feel like Being-experience but not all are truly so. And yet, we dare not neglect the clear hints that, sometimes at least, greater perspicuity and greater efficiency of cognition can be found in healthier people and in healthier moments, i.e., some peak-experiences *are* B-experiences. I once suggested the principle that if self-actualizing people can and do perceive reality more efficiently, fully and with less motivational contamination than we others do, then we may possibly use them as biological assays. Through *their* greater sensitivity and perception, we may get a better report of what reality is like, than through our own eyes, just as canaries can be used to detect gas in mines before less sensitive creatures can.

As a second string to this same bow, we may use ourselves in our most perceptive moments, in our peak-experiences, when, for the moment, *we* are self-actualizing, to give us a report of the nature of reality that is truer than we can ordinarily manage.

F. The After-Effects of Peak Experiences

Completely separable from the question of the external validity of cognition in the various peak experiences, is that of the after-effects upon the person of these experiences which in still another sense, may be said to validate the experience. I have no controlled research data to present to you. I have only the unanimous agreement of my subjects that there *were* such effects, my own conviction that there were, and the complete agreement of all the writers on creativeness, love, insight, mystic experience, and aesthetic experience. On these grounds I feel justified in making at least the following affirmations or propositions, all of which are testable.

1. Peak experiences have some therapeutic effects, in the strict sense of removing symptoms. I have at least two reports—one from a psychologist, one from an anthropologist—of mystic or oceanic experiences so profound as to remove neurotic symptoms forever after. Such conversion experiences are of course plentifully recorded in human history but so far as I know have never received the attention of psychologists or psychiatrists.

2. They change the person's view of himself in a healthy direction.

3. They change his view of other people and his relations to them in many ways.

4. They change more or less permanently his view of the world, or of aspects or parts of it.

5. They release him for greater creativity, spontaneity, expressiveness, idiosyncracy.

6. He remembers the experience as a very important and desirable happening and seeks to repeat it.

7. The person is more apt to feel that life in general is worth while, even if it is usually drab, pedestrian, painful or ungratifying, since beauty and goodness and excitement and honesty and truth and meaningfulness have been demonstrated to him to exist.

Many other effects could be reported that are ad hoc and idiosyncratic, depending on the particular person, and his particular problems which he considers to be solved or seen in a new light as the result of his experience.

I think that these after effects can *all* be generalized and a feeling for them communicated if the peak experience could be likened to a visit to a personally defined Heaven from which the person then returns to earth. Desirable after-effects of such an experience, some universal and some individual, are then seen to be practically inevitable.

And may I also emphasize that such after-effects of aesthetic experience, creative experience, love experience, mystic experience, insight experience, and other peak experiences are taken for granted and commonly expected by artists and art educators, by creative teachers, by religious and philosophical theorists, by loving husbands, mothers, and therapists. As a matter of fact, they are a commonplace to all but psychologists.

References

1. Allport, G. Becoming. New Haven: Yale Univ. Press, 1955.
2. Balint, M. Primary Love and Psychoanalytic Technique. New York: Liveright, 1953.
3. Bucke, R. Cosmic Consciousness. New York: Dutton, 1923.
4. Bühler, C. Maturation and Motivation. *Dialectica*, 1951, **5**, 312-361.
5. Dewey, J. Theory of Valuation. Int. *Encycl. Unified Science*. Chicago: Univ. Chicago Press, 1939.
6. Ghiselin, B., (*Ed.*). The Creative Process. Berkeley: Univ. California Press, 1952.
7. Goldstein, K. The Organism. New York: American Book, 1939.
8. Hartman, R. (Forthcoming book.)
9. Huxley, A. The Perennial Philosophy. New York: Harper, 1944.
10. ———. Heaven and Hell. New York: Harper, 1955.
11. Klee, J. The Absolute and the Relative. (Unpublished.)
12. Kris, E. Psychoanalytic Explorations in Art. New York: International Univ. Press, 1952.
13. Lee, D. *In* C. Moustakas (*Ed.*). *The Self*. New York: Harper, 1956.
14. Maslow, A. H. Motivation and Personality. New York: Harper, 1954.
15. Maslow, A. H. Deficiency motivation and growth motivation. *In* R. M. Jones (*Ed.*). *Nebraska Symposium on Motivation 1955*. Lincoln, Nebraska: Univ. Nebraska Press, 1955.
16. Northrop, F. The Meeting of East and West. New York: Macmillan, 1946.
17. Perls, F., Hefferline, R., & Goodman, P. Gestalt Therapy. New York: Julian Press, 1951.
18. Rogers, C. *In* C. Moustakas (*Ed.*). *The Self*. New York: Harper, 1956.
19. Schachtel, E. On Memory and Childhood Amnesia. *Psychiatry*, 1947, **10**, 1-26.
20. Scheler, M. The Nature of Sympathy. London: Routledge & Kegan Paul, 1954.
21. Schwartz, F. (Forthcoming book.)
22. Shlien, J. (Unpublished memoranda.)
23. Sorokin, P. The Ways and Power of Love. Boston: Beacon Press, 1954.
24. Tillich, P. The Courage To Be. New Haven: Yale Univ. Press, 1952.
25. Vivanti, L. A Philosophy of Potentiality. London: Routledge & Kegan Paul, 1955.
26. Werner, H. Comparative Psychology of Mental Development. New York: Harper, 1940.

Department of Psychology
Brandeis University
Waltham 54, Massachusetts

The Journal of Genetic Psychology, 1959, **94**, 67-76.

AN EXPERIMENTAL INVESTIGATION OF THE RELATIVE SPEEDS OF LEFT AND RIGHT-HANDED WRITERS*

Kingston College Clinic, Kingston Upon Hull, England

A. C. SMITH AND G. F. REED[1]

A. INTRODUCTION

In his book, *The Backward Child*, Sir Cyril Burt (4) remarked that the attitude of teachers towards left-handed children had changed subsequent to the First World War. Before that it had been customary to induce or compel the child showing an initial preference for the left hand at least to attempt to learn to write with his right hand. The development of a more permissive policy was attributable in part to the diffusion of theories which associated stammering with interference with a tendency to use the left hand (12, 15), in part, perhaps, to the simplifying of calligraphy and the spread of more liberal concepts of education, particularly at the infant stage. Burt suggested, indeed, that the apparent increase in the number of left-handed children entering the junior schools in the early 1930's was merely a reflection of the change in outlook among infant teachers (4). The many studies in this field have emphasized, among other things, the importance of making an early and accurate diagnosis of the strength of an initial left-handed tendency (11, 14, 16). At the same time, practical considerations have led, from as early as 1917, to investigations into the development of effective writing habits in the left-handed (1, 6, 7). It has been stressed, and stressed repeatedly, that writing is peculiarly difficult for left-handed people "—all systems of writing are based on the assumption that the writer will use his right hand" (5).

Some authorities suggest that the modes of movement, or movement and perception in combination, which handwriting involves are suited to right-handed people, but are difficult to adapt to the needs of the left-handed. Orton, for instance, observes that in writing "the left-handed child has to acquire an orientation and progress opposite to that which would be most facile for him" (12). Similarly Baker asserts that "the left-handed child

---

*Received in the Editorial Office on October 29, 1957.

[1] The authors are indebted to Mr. D. Speakman for his advice in the planning of a pilot study, and to Mr. C. A. Watcham and his Staff who made the pilot study possible; to Mr. A. Royse of Hull University for his statistical advice; to Dr. J. Mackay for his advice regarding the neural implications; and to the Staffs and schools who cooperated in the investigation.

naturally wishes to write from right to left . . . which would result in what is known as mirror-writing. . . . Since this style of writing is not ordinarily legible, the left-handed child resorts to awkward and inefficient methods of writing with the left hand" (2).

It appears probable that other contributory factors are to be sought in the mechanical difficulties of writing which vary in degree and in kind according to the hand preferred. Thus, where the right-handed writer draws the point of his pen or pencil across the page, the left-handed writer (except in comparatively rare cases) has to push it. Moreover, unless he learns to incline his paper towards the diagonal top-right, bottom-left (where the right-hander inclines it top-left, bottom-right), the left-handed writer tends to mask the first letters of a half-written word and, as he graduates to the use of pen and ink, risks smudging these first letters with his writing hand. These difficulties are further complicated by considerations of eye-dominance (13) and of the writing movement's inducing an increasing flexion of the left-hander's writing arm as against the increasing extension of the arm of the right-handed writer.

There is, then, little doubt that the left-handed child must find it harder to learn to write than the right-hander. But it is of greater interest to know whether this initial handicap persists throughout the left-hander's school career. Such a conclusion has been drawn repeatedly. And where effective *speed* of writing is concerned it has been supported by actual comparisons of left- and right-handed children.

Hildreth (1947) states: "The left-handed child writes with about four-fifths the speed of the normal right-handed child" (10). Burt (1937) reports: "Nearly always, whether old or young, the left-handed writer pens his lines more slowly. In actual tests for quickness I find that on the average the left-handed reach only four-fifths the speed of the right-handed" (4). And Garrison (1950) says of the left-handed writer: "Not only does he write with considerable tension, but his writing is significantly slower than that of other children of his age and grade level" (8).

B. Investigation

The test battery to be discussed here was designed with three objects in view: (*a*) To determine whether such differences in speed of hand-writing can be demonstrated. And, supposing such difference to be demonstrable, to determine (*b*) Whether they apply also to tasks which are less skilled or less familiar. (*c*) Presuming that the direction of movement required in hand-writing with its concomitant muscular, perceptual, and mechanical difficulties is a major factor contributory to such differences, to estimate the effect on these differences of a reversal in direction.

1. *Test Battery and Procedure*

The battery consisted of three tests: a simple repetitive writing test, and Tracing and "Ladders" tests. The two latter were designed to measure unpractised skills. The subjects were required to use pencils only.

2. *Simple Repetitive Writing Test*

To determine speed in a simple handwriting task, the children were given sheets of ordinary school writing paper lined at 8 mm. intervals. At the head of each sheet the words "cats and dogs" were written. Each child was required to write these words in print or in script as he chose as neatly and as often as he could with his preferred and non-preferred hands. The sequence was varied so that some children began with the preferred and some with the non-preferred hand. Sixty seconds were allowed for each part of this test.

3. *Tracing Test*

The subject was required to trace four lines along the long axis of a tracing board. These lines incorporated horizontal, vertical, and diagonal segments. The total length of each line amounted to 620 mm. in the proportions: horizontals 3, verticals 3, diagonals 4.

Direction of tracing and use of preferred and non-preferred hands were alternated in such a way as to provide measures of speed and accuracy in either direction with either hand. All subjects began the test with the long axis of the tracing board parallel to the near edge of their desks, though they were not prevented from making minor changes in the orientation of the board during the test. Use of preferred and non-preferred hands and directions of tracing were varied within the groups tested with the object of distributing the effects of training.

Thirty seconds were allowed for the tracing of each line and times noted for those children who completed a tracing within the time-limit. An approximate measurement of accuracy was made by deducting from the total score of millimetres traced the length of any segment in which the tracing deviated from the copy line for half its length or more.

4. *"Ladders" Test*

Each child was given a sheet of ordinary school paper lined at 8 mm. intervals. The paper was placed on the desk with the lines running at right angles to the desk's near edge, though, as in the tracing test, the children were not prevented from making small alterations to its position while they were working. The children were told that the two extreme lines were the sides of a ladder and asked to fill in the "steps," i.e., rungs, with horizontal pencil-strokes, each stroke touching, but not crossing, the vertical lines. The time allowed for each part of this test was 45 seconds and accuracy was assessed by subtracting from the total the number of "steps" which either failed to meet the verticals or crossed them on either side to a distance of 4 mm. or more. The sequence was varied so that some children began with the preferred hand and some with the non-preferred hand.

The subjects were tested in pairs for each test and were seated in such a way as to prevent their seeing each other's work. Facilities varied somewhat between schools, however, and it was not possible to provide strictly uniform conditions for testing.

5. *Subjects*

One hundred forty children were tested in all, 70 girls and 70 boys, aged 8-14 drawn from 6 schools. Four of these (ABCD) were primary schools, one (E) was a secondary modern school, and the sixth (F) comprised the Senior boys' and girls' departments of an unreorganized school. The numbers and compositions of these groups are set out in Table 1.

TABLE 1
DISTRIBUTION OF SUBJECTS BY SCHOOLS, SEX, AND HANDEDNESS

Schools		A	B	C	D	E	F	Totals	
Left-handed	Boys	5	5	4	6	6	10	36	70
	Girls	5	5	6	4	4	10	34	
Right-handed	Boys	7	5	4	6	6	10	38	70
	Girls	3	5	6	4	4	10	32	

The groups used for the investigation are believed to represent a fair cross-section of the local school population excluding the extremes of ability.

To avoid problems of definition the children forming the "left-handed" groups were merely those who use their left hands consistently for writing. No attempt was made to establish degree of laterality. These left-handed writers included the majority available in the schools used. In those cases where there was a choice of subjects, the desired number of children was taken in alphabetic order of surnames. The same sampling method was used for selecting the subjects composing the right-handed groups but these were roughly paired with the left-handers to the extent that they were drawn from the same schools and age-groups and, with the exception of one school, by sex.

C. RESULTS AND CONCLUSIONS

While the results show a tendency for the right-handed groups to write more rapidly than the left-handed where both sets of children used the preferred hand, the difference in writing speeds is not statistically significant either for the total of group differences or for the differences between the groups considered severally, school by school (Tables 1-4). These results, then, differ somewhat from the findings of Burt (4) and Hildreth (10).

Two possible reasons for the disparity of these results with those of earlier investigators may be suggested here. In the first place, the nature of the material chosen for testing speed of writing may not have favoured the emergence of differences of the kind expected. It may also be that over a prolonged writing period, the left-handers would more rapidly show the effects of fatigue. In the second place, the populations tested, despite the care taken to make them representative of the local school population, may have differed in marked degree from those tested in the earlier studies. In particular, it is possible that changes in educational policy of the sort indicated by Burt (4) have altered the composition of the population of children who write with their left hands. For it is likely that, as increasing numbers of children have been allowed to write with their left hands, the population of left-handed writers will have included more and more children whose preference for that hand is relatively weak. In contrast to this, the population of left-handed writers in the early years of the century probably comprised an irreducible minority of children with very strong left-handed preference determined by constitution or pronounced temperamental bias. It is also possible that the development in recent years of methods for teaching the left-handed child to write with that hand and the adoption of a more tolerant attitude on the part of authority have combined to make the left-hander better able to overcome the initial difficulties of learning to write.

No statistically significant differences in speed of working were found between the left- and right-handed group in the "Ladders" and Tracing tests when preferred hands were used. Nor were statistically significant differences found for these tests when Corrected Scores (i.e., Scores less Errors) were assessed.

As noted above, it has been said that for the left-handed writer movement across the page from left to right is less facile or natural than movement in the opposite direction. It may also be supposed that the progressive flexion of the left arm as it makes such a movement is less suited to skilled performance than the progressive extension of the right arm moving in the same direction. With a view to arriving at some evaluation of these assumptions performance

TABLE 2
TOTAL MEANS (N = 140)

Test	Group	Hands used			
		Preferred		Non-preferred	
		Uncorrected	Corrected	Uncorrected	Corrected
Tracings right-left	Right-handers	57	52	47	37
	Left-handers	55	47	47	42
Tracings left-right					
Score = amount traced to nearest cm.	R/H	62	53	50	42
	L/H	58	57	50	44
Ladders					
Score = number of strokes made	R/H	38	23	27	13
	L/H	36	22	27	16
Writing					
Score = number of letters written	R/H	78	72	33	25
	L/H	70	64	33	26

with the non-preferred hand was measured for both sets of groups on the writing test. Here the left-handed group were making tensor movements (from left to right with the right hand) and the right-handed groups were making movements of flexion in the same direction with the left hand. No statistically significant differences were observed.

TABLE 3
Analyses of Variance Raw Scores

	Preferred hands			Non-preferred hands		
S. of V.	d.f.	S. of S.	V	d.f.	S. of S.	V
Tracings right-left						
Hands	1	0.2	0.2	1	0	0
Sexes	1	0.2	0.2	1	0	0
Schools	5	4.9	0.98	5	2.4	0.48
Residual	16	8.1	0.50	16	9.5	0.59
Total	23	13.4	/	23	11.9	/
		F = 0.4			F = 0	
Tracings left-right						
Hands	1	0.66	0.66	1	0	0
Sexes	1	0.66	0.66	1	0.7	0.7
Schools	5	2.0	0.40	5	2.9	0.58
Residual	16	6.68	0.41	16	5.8	0.36
Total	23	10.00	/	23	9.4	/
		F = 1.6			F = 0	
Ladders						
Hands	1	0.17	0.17	1	0	0
Sexes	1	0.17	0.17	1	0	0
Schools	5	3.34	0.6y	5	1.5	0.30
Residual	16	4.16	0.26	16	3.0	0.18
Total	23	7.84	/	23	4.5	/
		F = 0.65			F = 0	
Writing						
Hands	1	2.67	2.67	1	0	0
Sexes	1	4.17	4.17	1	2.6	2.6
Schools	5	29.34	5.87	5	4.0	0.8
Residual	16	13.16	0.82	16	1.9	0.11
Total	23	49.34	/	23	8.5	/
		F = 3.25			F = 0	

The Tracing test results were scored both in terms of hands used and in terms of direction of movement. If outward movements involving extension of the tracing arm are, in fact, easier to make than inward ones involving flexion, higher scores should be made on the outward tracings than on the inward ones. While none of the differences observed were statistically significant, the general trend of the scores does not appear to support this assumption. In particular, it was noticeable that the corrected scores of the left-handed group tracing with their left hands in the left to right direction (i.e., making flexor movements) were somewhat better than those of the right-handed group using their right hands in the same direction (i.e., making tensor movements) and considerably better than their own left-handed scores in the right to left direction which also involved a tensor movement. In the "Ladders" test, however, in which the children were free to make their horizontal strokes in whichever direction they chose, a majority (77 per cent) were observed to make their strokes outward both with preferred and non-preferred hands.

TABLE 4
ANALYSES OF VARIANCE CORRECTED SCORES

S. of V.	Preferred hands d.f.	 S. of S.	 V	Non-preferred hands d.f.	 S. of S.	 V
Tracings right-left						
Hands	1	1.1	1.1	1	1.0	1.0
Sexes	1	0.1	0.1	1	0	0
Schools	5	10.4	2.08	5	6.7	1.34
Residual	16	7.1	0.44	16	7.3	0.45
Total	23	18.7	/	23	15.0	/
		F = 2.5			F = 2.22	
Tracings left-right						
Hands	1	0.7	0.7	1	0.1	0.1
Sexes	1	0.2	0.2	1	0.1	0.1
Schools	5	3.4	0.68	5	10.3	2.06
Residual	16	5.6	0.35	16	5.3	0.33
Total	23	9.9	/	23	15.8	/
		F = 2.0			F = 0.30	
Ladders						
Hands	1	0.1	0.1	1	37.4	37.4
Sexes	1	0.4	0.4	1	20.1	20.1
Schools	5	3.4	0.68	5	67.3	13.46
Residual	16	2.8	0.17	16	146.5	9.15
Total	23	6.7	/	23	271.3	/
		F = 0.58			F = 4.08	
Writing						
Hands	1	2.7	2.7	1	4.2	4.2
Sexes	1	4.2	4.2	1	121.5	121.5
Schools	5	24.4	4.88	5	226.4	45.28
Residual	16	10.1	0.63	16	149.8	9.36
Total	23	41.4	/	23	501.9	
		F = 4.28			F = 0.44	

It has also been suggested that the population of left-handed writers includes a substantial proportion of people in whom the functional differentiation of preferred and non-preferred hands is not so strongly developed as that of the bulk of the right-handed population. Supposing this to be true, one would expect the difference in performance between the preferred and non-preferred hands of the left-handed group to be rather less marked than that of the right-handed groups, particularly on the "Ladders" and Tracing test which measured unpractised skills. While no statistically significant differences were found, there was a general tendency in support of this view on all tests except left-to-right tracing (corrected scores). In this connexion it is interesting to note that the means of the corrected scores for the "Ladders" test when preferred hands were used were very slightly higher for the right-handed than for the left-handed groups, but when corrected scores for non-preferred hands were considered, the left-handed groups were found to be superior to .the right-handed to a degree which approached significance. A similar tendency has already been noted for the corrected scores in the Tracing Test and is reflected in a preliminary qualitative survey of the Writing Test which will be considered in detail in another paper.

In general, these findings seem to indicate a relatively higher degree of skill for the non-preferred hands of the left-handed children and hence a lower ratio between skilled performance in either hand than obtains for the right-handed children.

The data lend themselves to one further conclusion which may have important bearings. For the analysis of variance indicates clearly that neither handedness nor sex has affected the results of our tests to anything like the same degree as the different schools from which the groups were drawn (it was established that this was not merely a function of the different age groups involved). The implication, at this stage of our investigation, would appear to be that environmental factors are playing a greater part than neural differences with their associated muscular, perceptual, and mechanical difficulties. To assess fully the significance of this finding would evidently entail longitudinal studies which examined not only individual differences but social factors and the attitudes and methods of the teachers involved.

The major practical outcome, however, is that the assumption that the left-handed writer's problems necessarily result in slower handwriting cannot be supported by the results of this study.[2]

REFERENCES

1. BALLARD, P. B. How to write with the left hand. *Sch. World,* 1917-1918, **19-20.**
2. BAKER, H. J. An Introduction to Exceptional Children. New York: Macmillan, 1953.
3. BEELEY, A. L. An experimental study in left-handedness, with practical suggestions for schoolroom tests. Univ. Chicago: Educ. Monog., 1918, No. 8.

---

[2] Since completion of the above the writers have had the opportunity of reading Dr. Margaret Clark's excellent and authoritative book, recently published. It was too late to incorporate reference to this work in our paper, so that it is gratifying to note that her findings regarding the speed of writing of 18 left-handed writers are generally in accordance with our own.

4. Burt, C. The Backward Child. London: U.L.P., 1937.
5. Cole, L. Psychology of the Elementary School Subjects. New York: Farrar & Rinehart, 1934.
5a. Clark, M. M. Left-Handedness: Laterality Characteristics and Their Educational Implications. London: U.L.P., 1957.
6. Condon, R. J. Advice with regard to the treatment of left-handed writers. *Ele. Sch. J.*, 1927, **27**, 490-.
7. Gardner, W. H. Left-Handed Writing Instruction Manual. Danville, Ill.: Interstate, 1945.
8. Garrison, K. C. The Psychology of Exceptional Children. New York: Roland Press, 1950.
9. Haefner, R. Educational Significance of Left-Handedness. Teach. Coll., Columbia Univ., 1929.
10. Hildreth, G. Learning the Three R's: A Modern Interpretation. Minneapolis, Minn.: Univ. Minnesota, 1947.
11. ———. The development and training of hand dominance: I-V. *J. Genet. Psychol.*, 1949, **75**, 197-275; 1950, **76**, 39-144.
12. Orton, S. T. Reading, Writing, and Speech Problems in Children. New York Norton, 1937.
13. Parsons, B. S. Left-Handedness: A New Interpretation. New York: Macmillan, 1924.
14. Scheidemann, *et al.* The Psychology of Exceptional Children. Boston: Houghton Mifflin, 1931.
15. Terman, L. M. Hygiene of the School Child. New York: Macmillan, 1914.
16. Trankell, A. The influence of the choice of writing hand on the handwriting. *Brit. J. Educ. Psychol.*, 1956, **26**, No. 2.
17. Van Riper, C. The quantitative measurement of laterality. *J. Exper. Psychol.*, 1935, **18**, 372-382.

Kingston College Clinic
48 Beverley Road
Kingston Upon Hull
England

The Journal of Genetic Psychology, 1959, **94**, 77-83.

EMOTIONALITY AND AGGRESSIVE BEHAVIOR IN THE MOUSE AS A FUNCTION OF INFANTILE EXPERIENCE*

Columbus Psychiatric Institute and Hospital, Columbus, Ohio

SEYMOUR LEVINE[1]

A. THE PROBLEM

Several recent experiments (8, 9) have shown that rats given some form of experience in infancy (shock or handling) were more emotionally stable in adulthood than either late-handled or non-handled *S*s. There is, however, little available evidence concerning the effects of infantile (pre-weaning) experience on adult behavior in the mouse.

Hall and Whiteman (6) exposed young mice, from four to seven days of age, to the stimulation of an intense sound. When 30 days old, the *S*s were tested for emotionality by measuring the frequency of defecation in an open field. *S*s given the early stimulation were less stable emotionally when so tested. Beach and Jaynes (1) point out the possibility, however, that the results may have been due to specific and direct conditioning to environmental cues, since both the stimulation in infancy and the later test for emotionality were carried out in the same environmental setting. In addition, these investigators did not include as a control a group of non-handled *S*s. Stanley and Monkman (10) report no difference in avoidance learning between mice shocked in infancy (8-11 days) and mice not shocked during this time. However, again a non-handled control group was omitted.

Therefore, the only comparisons available in the mouse are between *S*s receiving differential treatments, rather than between treated and untreated *S*s.

One of the present studies investigated the effects of handling in infancy on emotionality measured on an elevated runway. The second experiment was designed to study the effects of infantile experience on aggressive behavior.

Hall and Klein (5) compared emotional and non-emotional rats for ag-

---

*Received in the Editorial Office on October 30, 1957.

[1] This investigation was supported by a research grant PHS M-1051 and M-1630 from The National Institute of Mental Health, Public Health Service. The mice for this study were made available by a research PHS M-123 from The National Institute of Mental Health.

The data for this paper were collected during the summer of 1956 at the Division of Behavior Studies of the Roscoe B. Jackson Memorial Laboratory. The writer wishes to express his appreciation to the staff of the Jackson Laboratory and especially to Drs. J. P. Scott and J. A. King for their assistance and support.

gressiveness and report that the non-emotional rats were significantly more aggressive than the more emotional *S*s. Generalizing from this finding and the finding that rats handled in infancy are more emotionally stable, it was predicted that mice handled in infancy would exhibit more aggressive behavior in adulthood as a result of the greater emotional stability achieved by the infantile experience.

B. Experiment 1

1. *Method*

a. Subjects: The mice used in this study were 24 C57 BL/10 female mice.

b. Apparatus: The apparatus consisted of an elevated runway 30″ × 1½″. The runway was 18″ from the floor and was enclosed on four sides to produce a homogeneous field. The runway was divided into 15 two-inch sections. The apparatus was illuminated from above by a 150 watt bulb approximately 15″ above the runway. The experimental room was lightproof.

c. Procedure: The *S*s were divided into two groups. Twelve *S*s were handled from Day 3 through Day 22. The remaining 12 *S*s were not handled in any manner until weaning on Day 22. Handling consisted of removing the *S* from the nest with forceps and placing *S* in a standard laboratory evaporating dish, 80 mm in diameter and 50 mm in height, for three minutes. At the end of this three-minute period, *S* was returned to the nest. This procedure was followed once daily until weaning (Day 22).

At weaning, the *S*s were placed in group cages (4-6 *S*s per cage) until Day 50, at which time the adult testing was initiated. Each *S* was given five trials of three minutes duration with approximately five minutes separating each trial. In each trial *S* was placed in the center section of the runway which was equidistant from both ends. Latency to leave the center section was recorded. Leaving this section was defined as entering the adjoining section with all four feet. In addition, defecation on the runway was recorded.

2. *Results*

No differences were observed between the groups in defecation. The mean number of trials on which defecation occurred was 4.25 for the handled group and 3.83 for the non-handled group.

However, the difference in latency to leave the center section on the runway was highly significant. The non-handled *S*s had significantly higher latencies than the handled *S*s.

Table 1 presents the analysis of variance of the latencies. Figure 1 presents the curves of the latency over the five trials.

TABLE 1

ANALYSIS OF VARIANCE OF LATENCY ON RUNWAY

Source	df	Mean square	F
Independent observations			
Between experience	1	5.77	13.11**
Between *S*s in same group	22	.44	
Total between *S*s	23		
Correlated observations			
Between trials	4	1.72	12.29**
Trial × experience	4	.27	1.93
Pooled *S*s × trials	88	.14	
Total within *S*s	96		
TOTAL	119		

**P = <.01.

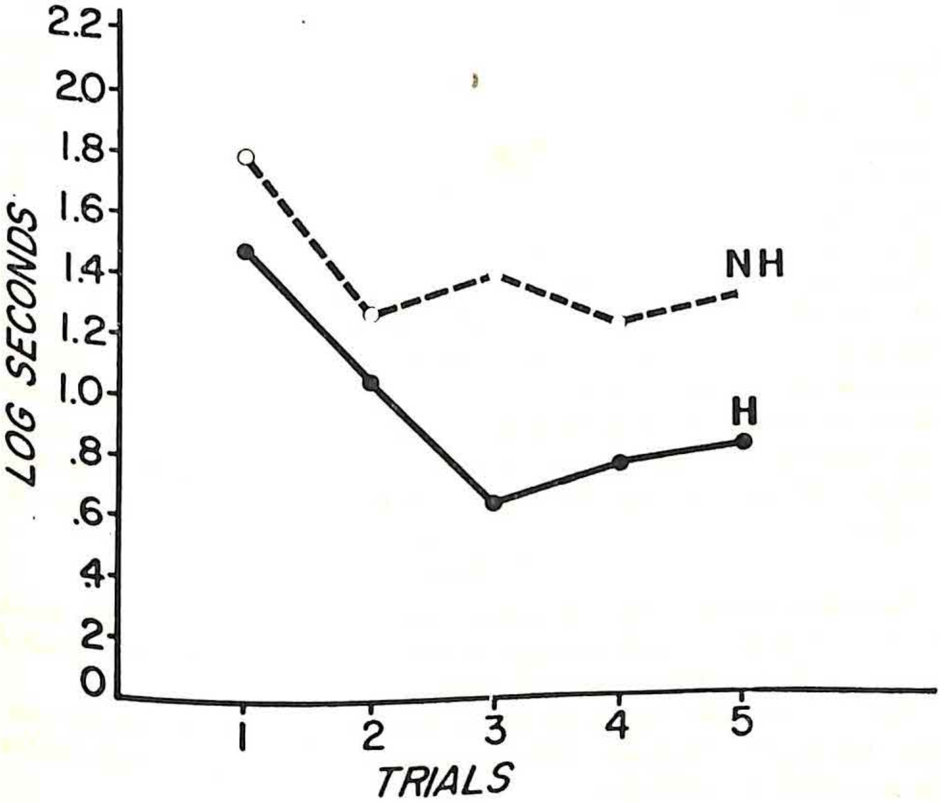

FIGURE 1

MEAN LATENCY FOR FIVE RUNWAY TRIALS: *H*, HANDLED IN INFANCY, *NH*, NON-HANDLED

C. EXPERIMENT 2

1. *Method*

a. Subjects: The mice used in this study were 36 C57 BL/10 males.

b. Apparatus: The *S*s were fought in boxes similar to those in which they had been housed. During the trials on which fighting was recorded, the partition separating the two 12″ × 6″ × 6″ living cages was removed, giving the *S*s free access to each other for that trial.

c. Procedure: Eighteen *S*s were handled from Day 2 through Day 22 in the manner identical to that described in the previous section of this paper. The remaining 18 *S*s received no handling. At weaning all *S*s were placed in individual cages.

Three days prior to testing, *S*s within each group were paired with other *S*s of their group in fighting boxes. No two *S*s from the same litter were paired. Each pair of mice was exposed to each other for seven days; one trial daily. Five minutes was the maximum time for each trial. The principle measure of aggressive behavior was the fighting response latency (4) which is defined as the elapsed time from the removal of the partition until a vigorous fight begins. Each pair was permitted to fight for five seconds, and was then separated by replacing the partition.

2. *Results*

Fighting response latencies for both groups on each of the seven trials is presented in Figure 2. Table 2 presents the analyses of variance for fighting response latencies.

Inspection of these data indicates that *S*s handled in infancy are more aggressive than non-handled *S*s. The non-handled group had significantly longer fighter response latencies than the handled group. This difference is not only reflected in the over-all latencies but further, an analysis of the differences on each trial indicates a significant difference beyond the .01 level on Trials 1, 5, 6; a difference beyond the .02 level on Trials 2, 3, and a difference at the .05 level on Trial 4. There was no significant difference on Trial 7.

An additional feature of these results is the difference in fighting on the first trial. In the handled group every one of the nine pairs of handled *S*s fought within the five minutes allotted for the trial, whereas only four of the nine pairs of non-handled *S*s fought. Because of the small *N*'s involved, Finney's (2) tables for marginal frequencies less than 15 were used to obtain the *P* value. The difference between the groups is significant at the .025 level of confidence.

TABLE 2
ANALYSIS OF VARIANCE OF FIGHTING RESPONSE LATENCY

Source	df	Mean square	F
Independent observations			
Between experience	1	36.46	14.70**
Between *S*s in same group	16	2.48	
Total between *S*s			
Correlated observations			
Between trials	6	23.57	38.18**
Trials × experience	6	.39	.58
Pooled *S*s × trials	96	.67	
TOTAL	125		

**P = <.01.

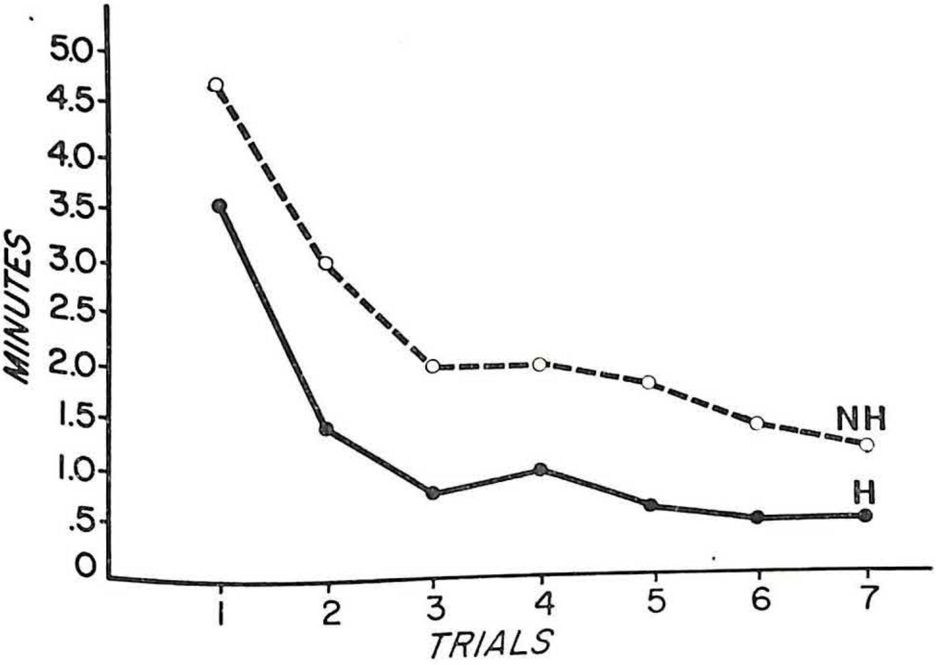

FIGURE 2
MEAN FIGHTING RESPONSE LATENCY

D. DISCUSSION

Recently, Willingham (11) isolated six factors related to emotionality in the mouse. Among these factors were elimination and freezing. Although there were no differences in elimination between the groups in this study, the non-handled group did show significantly longer emergence latencies (freezing). These results appear to be consistent with the previous data obtained, using rats and dogs (3), and are believed to further highlight the importance of varied infantile experience on the development of emotional stability in mammals.

The differences between the handled and the non-handled groups in aggressive behavior also appears to be consistent with the emotional stability hypothesis. King and Gurney (7), in attempting to explain why mice raised in groups, until 45 days of age, were more aggressive than *S*s raised in isolation when tested at 100 days of age, offered two explanations. The first explanation emphasized the learning of aggression through the competition occurring in the group cages. The other explanation states that "the mice merely learn what another mouse is, so when confronted with another mouse for the first time, it is not entirely strange." Thus these authors suggest that the innate tendency for aggressiveness is inhibited in proportion to the strangeness of the situation. The present study tends to support the latter explanation. It has previously been hypothesized that non-handled *S*s have a lower threshold for emotional susceptibility, and that novel stimuli result in a heightened emotional arousal sufficiently intense to interfere with adaptive behavior.

In the present situation, which was identical to that of King and Gurney, one-half of the fighting box is new for each mouse, and since these *S*s have been isolated since weaning, each member of the pair is strange to the other member. Further, the lifting of the partition is an additional component of novelty in the situation. Observation of the behavior of the non-handled *S*s adds additional support for the novelty hypothesis. Upon removal of the partition, these *S*s tended to crouch in the corner of the box for some time before engaging in exploration of the box, and of the other mouse. This was particularly true on the first trial, and is reflected in the significant differences between the groups in the number of *S*s which fought on the first trial.

The similarity between the emergence latencies and fighting response latencies should be noted. It seems possible that we are dealing with the same behavior, namely, freezing, in two distinct situations. Further, the heightened emotionality as manifested by freezing in the non-handled animals appears to affect many kinds of behavior.

E. SUMMARY

Sixty C57 BL/10 mice were handled in infancy to determine the effects of infantile experience on emotionality and aggressive behavior later in life. Control groups received no handling.

At 50 days of age a group of 12 handled female mice were compared with a group of non-handled mice in the response to an elevated runway. In addition, nine pairs of handled mice were compared with nine pairs of non-handled male mice on latency to initiate a fight.

The results on the runway indicated that handled *S*s had significantly shorter latencies to initiate movement on the runway than the non-handled controls. Further, the handled male mice had significantly shorter fighting response latencies than the non-handled *S*s.

The results were interpreted as indicating that the non-handled *S*s have a lower threshold for emotional susceptibility than do the handled *S*s and that novel stimuli result in a heightened emotional arousal which is manifested by freezing in non-handled mice.

REFERENCES

1. BEACH, F. A., & JAYNES, J. Effects of early experience upon the behavior of animals. *Psychol. Bull.,* 1954, **51**, 239-263.
2. FINNEY, D. J. The Fisher-Yates test of significance in 2 × 2 contingency tables. *Biometrika,* 1948, **38**, 145-156.
3. Fisher, A. E. Differential early treatment as an experimental variable. *Amer. Psychol.,* 1956, **11**, 427.
4. FREDRICKSON, E. Competition: The effects of infantile experience upon adult behavior. *J. Abn. & Soc. Psychol.,* 1951, **46**, 406-409.
5. HALL, C. S., & KLEIN, S. J. Individual differences in aggressiveness in rats. *J. Comp. Psychol.,* 1942, **33**, 371-383.
6. HALL, C. S., & WHITEMAN, P. H. The effects of infantile stimulation upon later emotional stability in the mouse. *J. Comp. & Physiol. Psychol.,* 1951, **44**, 61-66.
7. KING, J. A., & GURNEY, N. Effect of early social experience on adult aggressive behavior in C57 BL/10 mice. *J. Comp. & Physiol. Psychol.,* 1954, **47**, 326-330.
8. LEVINE, S., CHEVALIER, J. A., & KORCHIN, S. J. The effects of early shock and handling on later avoidance learning. *J. Pers.,* 1956, **24**, 475-493.
9. ———. A further study of infantile handling and avoidance learning. *J. Pers.,* 1956, **25**, 70-80.
10. STANLEY, W. C., & MONKMAN, J. A. A test for specific and general behavioral effects of infantile stimulation with shock in the mouse. *J. Abn. & Soc. Psychol.,* 1956, **53**, 19-22.
11. WILLINGHAM, W. W. The organization of emotional behavior in mice. *J. Comp. & Physiol. Psychol.,* 1956, **49**, 345-353.

Columbus Psychiatric Institute and Hospital
University Health Center
Columbus 10, Ohio

The Journal of Genetic Psychology, 1959, **94**, 85-91.

ANXIETY AND ACCURACY OF PERCEPTION OF SOCIOMETRIC STATUS AMONG INSTITUTIONALIZED DELINQUENT BOYS*

New York State Training School for Boys

RICHARD D. TRENT

A. PURPOSE

Recent studies in sociempathy have reported that there is no relationship between the accuracy of perception of own status and accuracy of perception of others' status (2, 16). This evidence appears contradictory to the self-other hypothesis which states that attitudes and appraisals of self are positively related to attitudes and appraisals of others (4, 10, 11, 12, 13, 14). On the basis of the self-other hypothesis and assuming that self-attitudes are relatively stable (5), we would expect: (*a*) the greater the accuracy of perception of self status, the greater the accuracy of perception of others; (*b*) if high anxiety is a characteristic of the unhealthy self as Horney (8) has suggested, we would expect that more anxious *S*s would tend to be less accurate in perceiving their own and others' sociometric status; and (*c*) those *S*s who over- or underestimate their own sociometric status should tend to over- or underestimate the sociometric status of others.

The purpose of this inquiry was to determine the interrelationships between accuracy of perception of own status, accuracy of perception of others' status, and anxiety among institutionalized delinquent boys. An individual's awareness of his own and others' sociometric status in a given group of which he is a member is called sociempathic ability (1, 2). Anxiety was operationally defined in terms of *S*s' performance on the Children's Form of the Manifest Anxiety Scale (6).

B. SUBJECTS

The *S*s consisted of 63 boys in two cottages at the New York State Training School for Boys, a residential institution for adjudicated delinquents. Thirty-one *S*s lived in one cottage unit and 32 *S*s in the second cottage. In the training school a group differs from the usual educational group in that there is constant change: new members are added, old members are lost because of the weekly intake and release procedures. These two groups were

---

*Received in the Editorial Office on November 5, 1957.

groups mainly in the sense of being together in the same living situation. The average length of cottage stay at the time of the study was 18.7 weeks with a range of two weeks to 11 months.

Thirty-three *S*s were Negro, 16 were white, and 21 were of Puerto Rican origin. Thirty-seven were Protestant, 26 were Catholic. The average intelligence of *S*s, as measured by the Wechsler-Bellevue, was 81 with a range from 57 to 104. All *S*s were from deprived socio-economic backgrounds and high delinquency areas of New York City. Sixteen boys were from homes where both the natural father and mother were present; 47 were from homes where either the natural father or the natural mother were absent from the home.

C. PROCEDURE

Anxiety was measured by the Children's Form of the Manifest Anxiety Scale which will hereafter be called the *CMAS*. This scale was an adaptation of the Taylor Scale of Manifest Anxiety (15) and was developed by Castaneda (6) and his associates at the Iowa Child Welfare Research Station. The scale consists of two types of items, 41 anxiety items and 11 lying items (*L*-scale) which were designed to measure the subjects' tendency to falsify responses. Each item was read to the subjects by this investigator. Instructions given to subjects included: "I am going to read 53 statements to you. Listen carefully. Put a circle around the word *Yes* if you think the statement is true about you. Put a circle around the word *No* if you think the statement is not true about you." The *CMAS* score was determined by summing the total number of *Yes* responses circled for the 41 anxiety items.

Odd *CMAS* item scores were correlated with even *CMAS* item scores to obtain the reliability of the instrument. The reliability coefficient, corrected by the Spearman-Brown formula, was .78 ($p < .01$). *CMAS* scores ranged from 5 to 37 with a mean and sigma of 20.71 and 7.08 respectively. It is interesting to note that the scores of this group were significantly higher than those of the sixth grade boys of Castaneda's *CMAS* standardization group. The mean lying score was 5.56 with a sigma of 2.31. The correlation between anxiety score and lying scores was only .08 which indicates no relationship between scales. This latter correlation was desirable for it suggests that falsification of responses could result randomly in a high, as well as a low, anxiety score.

Accuracy of perception of own and others status was determined by direct questioning. Each boy was asked to rate each other boy in his group on a five-point scale indicating: (*a*) the way the boy felt about each other boy, (*b*) the way the boy thought each other boy felt about him, and (*c*) his prediction of the popularity of each boy in the group including himself. The specific procedure used has been fully enumerated in a previous study (16).

Accuracy of perception of one's own sociometric status was expressed as a total discrepancy score for each boy, representing the sum of the deviation of how others would rate him from the actual ratings given him by these boys. Self-discrepancy scores ranged from 21 to 66 with a mean and sigma of 38.37 and 10.2 respectively. To determine the reliability of self-accuracy scores, odd discrepancy scores were correlated with even discrepancy scores. A reliability coefficient, corrected by the Spearman-Brown formula, of .68 ($p < .01$) was found, indicating consistent performance of boys.

Accuracy of perception of others' sociometric status was expressed by the total discrepancy between all individual predictions made by a subject of the sociometric status of other and the actual sociometric status scores of the latter. A *S*'s sociometric status was ascertained by taking the mean of all the sociometric ratings he received from others. Accuracy of perception of others' scores ranged from 10-51 with a mean and sigma of 27.58 and 7.9 respectively. The corrected reliability correlation for the accuracy of prediction of others' status was .76 ($p < .01$) and was determined by the split-half method.

To answer two of the three predictions made, a matrix of intercorrelations between anxiety, accuracy of predicting self and others' status, actual social status, and length of cottage stay was developed. Sociometric status and length of cottage stay were included for inquiries have noted the relationship of these variables with accuracy of perception of self and others' status (2, 8, 16). The matrix of intercorrelations is shown in Table 1. There was no significant relationship between accuracy of perception of own status and accuracy of perception of others' status. Anxiety was negatively related

TABLE 1

A MATRIX OF CORRELATIONS BETWEEN ANXIETY AND FOUR VARIABLES

		Correlations				
Variables	No.	A	B	C	D	E
Anxiety	A		−.27*	−.11	−.31*	.08
Accuracy of perception of own status	B			.07	.47*	.04
Accuracy of perception of others' status	C				.22	.19
Social status	D					.11
Length of stay in the cottage	E					

*A correlation of .254 was significant at the .05 level of probability.

($p < .05$) to accuracy of perception of own status but unrelated to accuracy of perception of others' status. The less anxious were significantly more accurate in perceiving own status than were the more anxious.

The less anxious were significantly more popular ($p < .05$) than the more anxious. The more accurate the perception of own status the greater the social status. Length of stay was unrelated to the four variables studied.

Did those *S*s who overestimated their own status tend to overestimate the status of others? To answer this question, each *S*'s estimate of his own status was compared with his own actual status. It was found that 31 boys overestimated their own status, 21 boys correctly estimated their status, and 11 boys underestimated their status. Within each of these three groups, each *S*'s estimate of each other boy was studied to determine if the *S*s overestimated, underestimated, or correctly estimated each others' status. In the first cottage ($N = 31$), each boy had 30 opportunities to over-, under- or correctly estimate others' status. In the second cottage ($N = 32$), each boy had 31 opportunities to over-, under-, or correctly estimate others' status.

The number and per cent of overestimations, correct estimations, and underestimations made by *S*s grouped according to overestimation, correct estimation, or underestimation of their own status is shown in Table 2. It should be noted that the greatest contributions to the large percentage differences in Table 2 came from the overestimation of self-status group who greatly overestimated others' status and from the self-underestimators who tended strongly to underestimate the status of others.

TABLE 2
ACCURACY OF ESTIMATIONS OF OWN STATUS AS COMPARED TO THE ACCURACY OF ESTIMATIONS OF OTHERS' STATUS

Accuracy of estimations of own status	Accuracy of estimations of others' status: Over-estimated others' status No.	%	Correctly estimated others' status No.	%	Under-estimated others' status No.	%	Total estimations No.
Overestimated own status (N = 31)	444	64	279	44	191	34	914
Correctly estimated own status (N = 21)	195	28	231	36	214	38	640
Underestimated own status (N = 11)	59	8	125	20	154	28	338
Total	698	100	635	100	559	100	1892

D. Discussion

The results reported above should be regarded as tentative and should be interpreted in terms of the number and kinds of *S*s used and the validity of the instrument employed.

What implications do these results have for the theory of self and its relation to perception? Apparently, there is no relationship between the accuracy of perception of self-status and the accuracy of perception of others' status. It may be that the ability to perceive own status and the ability to perceive the status of others are separate abilities which involve distinctly different perceptual processes. Or, it may be that the real relationship between perception of own and others' status is contaminated by pervasive influences such as the desire of *S*s to react to others in a manner they considered appropriate within the particular living situation; or such as the interaction of individual estimates with reference group estimates. Whether or not the above suppositions are valid is a matter for further research.

High anxiety was inversely related to accuracy of perception of own status but unrelated to accuracy of perception of others' status. This finding seems consistent with the principle that an individual's own feelings and emotions markedly influence his perception of himself. Apparently, the *S*s of this study could appraise the status of others with more objectivity than they could appraise their own status.

The subjects of the study, who overestimated their own status, did demonstrate a significant trend to overestimate the status of others. This finding seems to support the self-other hypotheses; however, we may attribute this finding to simpler factors including: (*a*) the "halo effect," the tendency to assign a high or low rating to all, including himself, for the characteristics being measured, (*b*) some unknown artifact of the procedure employed, (*c*) some factor or factors relevant to the type of sample employed.

E. Summary

The major purpose of this study was to determine the interrelationships between accuracy of perception of own sociometric status, accuracy of perception of the status of others, and anxiety. Accuracy of perception of own status was measured by discrepancy scores which represented the sum of the deviations of how others would rate the individual on a five-point acceptance-rejection scale as compared to the individual's actual rating. Accuracy of perception of others' status was determined by a discrepancy score which was the sum of deviations of how the individual rated the sociometric status

of others, on a five-point scale, as compared to the other individual's actual rating. Anxiety was measured by the Children's Form of the Manifest Anxiety Scale. The *S*s consisted of 63 boys who lived in two cottage groups at a residential training school for delinquent boys.

The major findings of this study may be stated as follows:

1. There was no significant relationship between accuracy of perception of own status and accuracy of perception of others' status.

2. Anxiety was negatively related to accuracy of perception of own status but unrelated to accuracy of perception of others' status.

3. The less anxious enjoyed significantly higher social status than the more anxious. The more accurate the perception of own status, the greater the social status.

4. Those *S*s who overestimated their own status tended to overestimate the status of others.

These results were discussed in terms of their implications for the self-other hypothesis and perceptual theory. It was suggested that perception of own status may involve a different perceptual process than perception of others' status. Whether or not these findings may be generalized to dissimilar populations is a matter for further research.

REFERENCES

1. AUSUBEL, D. P., *et al.* A preliminary study of developmental trend in sociempathy: Accuracy of perception of own and others' sociometric status. *Child Devel.*, 1952, **23**, 111-128.
2. AUSUBEL, D. P., & SCHIFF, H. M. Some intrapersonal and interpersonal determinants of individual differences in sociempathic ability among adolescents. *J. Soc. Psychol.*, 1955, **41**, 39-56.
3. AUSUBEL, D. P. Sociempathy as a function of sociometric status in an adolescent group. *Hum. Rel.*, 1955, **8**, 75-84.
4. BERGER, E. M. The relationship between expressed acceptance of self and expressed acceptance of others. *J. Abn. & Soc. Psychol.*, 1952, **47**, 778-782.
5. BROWNFAIN, J. J. Stability of the self-concept as a dimension of personality. *J. Abn. & Soc. Psychol.*, 1952, **47**, 597-606.
6. CASTANEDA, A., *et al.* The children's form of the manifest anxiety scale. *Child Devel.*, 1956, **27**, 317-332.
7. GRONLUND, N. E. Sociometric status and sociometric perception. *Sociometry*, 1955, **18**, 122-128.
8. HORNEY, K. On feeling abused. *Amer. J. Psychoanal.*, 1951, **11**, 5-11.
9. MCCANDLESS, B. R., *et al.* Anxiety in children and social status. *Child Devel.* (In press.)
10. MCINTYRE, C. J. Acceptance of others and its relation to acceptance of self and others. *J. Abn. & Soc. Psychol.*, 1952, **47**, 624-625.
11. ROGERS, C. R. A coordinated research in psychotherapy. *J. Consult. Psychol.*, 1949, **13**, 149-153.

12. Scheerer, E. T. An analysis of the relationship between acceptance of and respect for self and acceptance of and respect for others in 10 counseling cases. *J. Consult. Psychol.*, 1949, **13**, 169-175.
13. Stock, D. An investigation into interrelations between the self-concept and feelings directed toward other persons and groups. *J. Consult. Psychol.*, 1949, **13**, 176-180.
14. Sullivan, H. S. Conceptions of modern psychiatry. Washington, D. C.: William Alanson White Inst. Psychiat., 1947.
15. Taylor, J. A. A personality scale of manifest anxiety. *J. Abn. & Soc. Psychol.*, 1953, **48**, 285-290.
16. Trent, R. D. Sociempathic ability in a group of institutionalized delinquent boys. *J. Genet. Psychol.* (in press).

Puerto Rico Institute of Psychiatry
P.O. Box 127
Bayamon, Puerto Rico

The Journal of Genetic Psychology, 1959, **94**, 93-100.

PARENT EVALUATION AS A DETERMINER OF IDEOLOGY*

Department of Psychology, San Jose State College

JOSEPH B. COOPER AND MARGARET A. BLAIR

A. INTRODUCTION AND THE PROBLEM

This exploratory study was broadly concerned with the problem of attitude genesis. Students of attitude genesis (8) have taken the view that attitudes are at least as much functions of the social milieu as functions of direct experiences the individual has with attitude objects. For example, Newcomb's (10) classical Bennington study pointed clearly to the over-all influence of reference groups as attitude determiners. And, in much the same vein, Gordon Allport (3) has stressed the importance of one's family in creating a condition which fosters the development of a particular attitude cluster.

Allport (3) has used the identification construct to describe the degree of closeness a child feels toward his parent, his family. This point has been expanded to throw light upon the child's acceptance or rejection of parental ideology. And beyond this, there are several evidences which indicate that a social ideological system is basic to the particular attitudes the individual develops and uses (1, 4).

Ideological system is used here as representing a complete, holistic view of self and world. It is not compartmentalized as separate from personality.

The problem may be stated as follows: Evidence supports the proposition that attitudes (especially prejudicial attitudes) are in great measure functions of ideology. Social psychologists generally hold that parents play a significant rôle in the formation of the individual's ideology. In view of this line of reasoning a question should then be asked. Does not the individual's evaluation of his parent (essentially, as being "good" or "bad") tend to determine whether or not the individual's ideological system will be similar to that parent's?

Of the many questions posed by this large, general problem, one major and two ancillary questions have been isolated, stated, and tested in this study.

The major question concerned the relationship between the individual's evaluation of his parent and the degree of ideological similarity between the individual and the parent. Stated as an hypothesis: An individual who indicates a high evaluation of a parent also indicates a close ideological re-

---

*Received in the Editorial Office on November 19, 1957.

semblance to that parent; and, an individual who indicates a low evaluation of a parent also indicates a wide discrepancy between his own ideology and that of his parent.

The next question—ancillary—concerned the psychoanalytic contention that parent preference is determined by opposite-sex status. The reasons for studying this suspected determining variable are several, but for present purposes there was one particular reason. Are there other than social learning factors which influence the individual's evaluation of the parent? If, according to the above hypothesis, the individual's regard for his parent has a determining effect upon his social ideology, then it seems important to inquire into variables suspected as being significant in determining parent evaluation. Stated as a proposition: The individual's parent preference is determined by an opposite-sex predilection.

The third question—ancillary—was directed to the prevalent notion that within our culture mothers are typically preferred to fathers. This question has been dealt with by both researchers (5, 7) and popular writers (13). Again, if the major hypothesis of this study is supported it then becomes significant to discover whether or not mothers are typically preferred to fathers. Stated as a proposition: The female parent is preferred by both sons and daughters to the male parent.

B. MATERIALS, SUBJECTS, AND PROCEDURES

1. *Materials*

Two questionnaire-type scales were used: a 150-item parent evaluation scale, and the 30 item *E-F* scale (6).

a. Parent Evaluation Scale. This scale (hereafter referred to as the *PE* Scale) was devised for the purpose of measuring attitude content and strength toward parents (9). Items developed for use in this scale emerged from dynamic conceptions of parent-child relationships.[1]

Item-statements described types of behavior or personality characteristics customarily looked upon as "good" or "bad."

Positive statements were designed to describe the following: acceptance, comradeship, helpfulness, consistency, reasonableness, assurances of love and affection, confidence in the child's good judgment, acceptance of the parent rôle, and wholesomeness of attitudes toward sex and dating.

Negative statements were designed to describe the following: favoritism for a sibling, rejection, neglect, punitiveness, over-protection, martyrdom, nagging, scolding, ridicule, coldness, lack of confidence in the child's authoritarian control, and unwholesomeness of attitudes toward sex and dating.

---

[1] This scale will be provided to readers upon request.

Also included were statements descriptive of personality characteristics thought to be important in the development of attitudes children develop towards their parents. Positive statements described such personality traits as cheerfulness, friendliness, hospitality, optimism, self-confidence, and a sense of humor. Negative statements were descriptive of moodiness, pessimism, irritability, stinginess, insecurity, quarrelsomeness, and unfriendliness.

Edited statements totaled 170. These were divided into four categories: mother-positive, 40; mother-negative, 45; father-positive, 43; father-negative, 42.

Students in four upper-division psychology classes judged the items for content and strength. Each class was given one of the four categories to judge. Judgments were made on an 11-point scale in accordance with the Thurstone (12) method of "equal-appearing intervals."

Each of the 170 items was evaluated by integrating three variables which resulted from the judgments: (*a*) number of ratings for the statement at each interval on the 11-point scale; (*b*) the modal interval position for the statement; and (*c*) calculation of "deviation scores" for judges' ratings placed at stations other than the modal station.

Deviation scores were calculated by $\frac{\Sigma D\ (f)}{N}$ in which D = distance, or number of stations from the mode; f = frequency of ratings at each station; and, N = number of judges rating the item.

In its final form the *PE* scale had a total of 150 statements, 75 applying to mothers and 75 to fathers. For each parent there were 38 positive and 37 negative statements.

The 25 positive and 25 negative statements having the smallest deviation scores were retained as scale items. Most of the statements with relatively large deviation scores were retained as "filler" items. In the final scale (for both the mother section and the father section) items were arranged randomly.

Scoring was based upon items checked—a one-point value being assigned to each. Weighted values (though available) were not used. Positive, negative, and algebraic scores were obtainable.

Two "filler" statements incorporated within the scale made it possible to establish two validity indexes. Although the procedure was not rigorous, it afforded some indication of the value of the instrument as a measure of parent evaluation. These statements were:

"I have felt closer to my mother than to my father."
"I have felt closer to my father than to my mother."

These statements are obviously indirect to the point of parent preference, but their use was based upon the assumption that information concerning "feeling of closeness" could be more easily and accurately elicited than direct information concerning the preference of one parent to the other. The declaration was tested against the algebraic *PE* scale score. The strong tendency was for those subjects who expressed a greater feeling of closeness for one parent than the other to score that parent relatively high on the *PE* scale. Chi square for parent preference declaration and *PE* scale scores was 22.72, $p < .0005$. The contingency coefficient was .39. This finding lends support to the view that the *PE* scale has sufficient validity for use in an exploratory study of this type.

Reliability of the *PE* scale was estimated by test-retest for the mother sub-scale and for the father sub-scale. Two samples were used. For the first sample: the mother sub-scale $N = 14$, and the father sub-scale $N = 15$. Also, for the first sample, *PE* scores were obtained by three separate methods: algebraic, plus, and minus scoring. Table 1 shows the sub-scale Spearman *r*'s for each scoring method.

TABLE 1
SPEARMAN *r*'S FOR SUB-SCALES

	Algebraic score	Plus score	Minus score
Mother	.9661	.9836	.7698
Father	.9197	.7774	.9956

For the second reliability sample the *N* was 41 for the mother sub-scale and 38 for the father sub-scale. Algebraic scoring was used. For the mother sub-scale the Spearman *r* was .9577, and for the father sub-scale it was .9691.

These coefficients suggest that, within the limitations of administration and sampling restriction, the scale has desirable reliability.

b. E-F Scale. The 30-item *E-F* scale, assembled by Gough (6), was used for the purpose of assessing ideology. The *E-F* scale was derived from the *E* (ethnocentrism) and *F* (implicit anti-democratic trends) scales to be found in *The Authoritarian Personality* (2). Scoring was conventional.

2. *Subjects*

The subjects used in this study were students at San Jose State College. From the original sample 179 were retained, an age restriction of 17 to 30 years being imposed. Of these, 112 were male and 67 were female.

Some few of the subjects were unable to complete all parts of the scales. A few subjects did not complete one of the three *E-F* scales administered. Also, a few subjects were unable to complete one or the other *PE* sub-scales. Such incompletions did not invalidate these subjects' other data, which could be used in appropriate analyses. This accounts, however, for *N* fluctuations from analysis to analysis.

3. *Procedures*

Scales were administered to subjects during regular class periods. Subjects were assured anonymity and promised pertinent information about results of the study.

Both the *PE* and *E-F* scales were accompanied by answer sheets. Sex and age, among other personal information items, were obtained.

Upon completion of the *PE* scale the *E-F* scale was administered. For the *E-F* scale, three answer sheets were provided. These were: (*a*) the subject's responses; (*b*) the way the subject believed his mother would respond; and (*c*) the way the subject believed his father would respond.

C. RESULTS

The hypothesis and the two ancillary propositions will be stated in turn. For each, test results will be provided.

The hypothesis: Subjects who evaluate their parents relatively high, as measured by the *PE* scale, tend to be in relatively close ideological agreement with their parents, as measured by the degree of difference between their responses to the *E-F* scale and the way they believe their parents would respond to the *E-F* scale.

The hypothesis was tested in three ways. Mother and father sub-scales were treated separately. First, algebraic *PE* scores were tested against *E-F* difference scores. Second, plus *PE* scores (i.e., number of positive statements checked) were tested against *E-F* difference scores. Third, the highest 20 *PE* (algebraic) scoring subjects were compared with the lowest 20 *PE* (algebraic) scoring subjects, with reference to *E-F* difference scores.

Results of the first test of the hypothesis were as follows. Subjects who scored (algebraically) mothers above the median were assigned to a high evaluation category, and subjects who scored mothers below the median were assigned to a low evaluation category. For subjects who scored mothers above the *PE* median the median *E-F* difference score was 18; and for those who evaluated mothers below the *PE* median the median *E-F* difference score was 25. This was as predicted by the hypothesis. The Mann-Whitney "*U*" Test (11) was applied. In this instance: $N = 159$, $z = 3.38$, $p < .0005$.

For subjects who scored fathers above the *PE* median the *E-F* difference score median was 11.5. For subjects who scored fathers below the median the *E-F* difference score median was 25. This, too, was as predicted by the hypothesis. The "*U*" test result was: $N = 147$, $z = 3.14$, $p < .0008$.

The second test of the first hypothesis resulted in the following. Subjects who scored mothers above the median (positive statements checked) had a median *E-F* difference score of 16. Subjects who scored mothers below the median had a median *E-F* difference score of 25. This was as predicted by the hypothesis. Subjects who scored fathers above the median had a median *E-F* difference score of 12.5. Subjects who scored fathers below the median had a median *E-F* difference score of 25. This, too, was in line with the hypothesis. The "*U*" test was again applied for mother and for father. For mother: $N = 148$, $z = 3.65$, $p < .00016$. For father: $N = 147$, $z = 2.96$, $p < .0015$. (*N* discrepancies between the first and second tests result from deletion of median *PE* scores.)

The third test also furnished evidence in support of the hypothesis. This test compared the 20 highest scoring *PE* subjects with the 20 lowest scoring *PE* subjects. *E-F* difference score medians were 7 and 45.5 respectively for mother; and, 11 and 29.5 respectively for fathers. The Kolmogorov-Smirnov test (11) was applied. For both mothers and fathers results were the same: $X^2 = 22.00$, $p < .001$.

Results from the three tests just described offer substantial support for the hypothesis. If parent evaluation can be operationally defined by *PE* scale measurement, and ideology can be operationally defined by *E-F* scale measurement, then there is evidently a strong tendency for persons who evaluate parents favorably to be close to parents ideologically.

The first proposition: Daughters tend to evaluate fathers higher than mothers as measured by *PE* scores; and, sons tend to evaluate mothers higher than fathers as measured by *PE* scores.

This ancillary, or "probing" proposition, designed to cast light upon the genesis of parent preference, was tested in two simple ways.

First, use was made of the two scale items which permitted subjects to declare parent preference by way of "feeling of closeness." Of the male subjects who made such declarations, 78 per cent indicated a mother preference; of the female subjects, 76 per cent indicated a mother preference. The difference is obviously non-significant.

Second, a division by sex was again made, but this time *PE* algebraic scores were used. Results were: 48 per cent of the males and 53 per cent of the females indicated mother preference. The difference is again obviously non-significant.

These simple findings are not to be viewed as refutations of the "Oedipal complex" construct. However, within the limitations of these data, no evidence is furnished in its support.

The second proposition: Female parents are preferred to male parents by both daughters and sons, as measured by the *PE* scale and parent preference declaration.

This proposition was tested in three ways.

First, subjects who completed both mother and father *PE* sub-scales were divided into two groups: mother preferred and father preferred. In this instance the two groups were equal in frequency: 83 and 83.

Second, subjects who *declared* a preference for mother or father ("feeling of closeness") were divided into two groups upon that basis. In all, 129 declared such a preference. Of these, 100 declared a mother preference and 29 declared a father preference. The binomial test (11) was applied: $z = -6.16$, $p < .00003$.

Third, the same 129 subjects (just above) were divided into two groups upon the basis of *PE* scores, with reference to father or mother preference. In this instance, 70 scored mother higher, and 59 scored father higher. Again, the binomial test was applied: $z = -.88$, $p < .1894$.

Results from these three tests of the second proposition are equivocal. The first and third tests indicate that when subjects score parents by way of the *PE* scale, fathers are preferred about as frequently as mothers. However, the second test, based on *declarations* of parent preference, show mothers to be preponderantly preferred. While no clear answer is furnished to this proposition it might be supposed that two different parent-perception variables were being assessed.

D. Summary

This was an exploratory study of the relationship between parent evaluation and parent-child ideology similarity. Basic to the study was the assumption that attitudes are determined in large measure by ideology. An hypothesis was stated to the effect that persons who evidence relatively high regard for their parents develop ideologies which are relatively similar to their parents'. In light of the plausibility of this hypothesis, two ancillary propositions were stated and tested. Both of these dealt with possible determiners of parental preference. One concerned the psychoanalytic thesis that parental predilection is determined by opposite-sex status. The other proposition concerned the frequently stated view that, within our culture, mothers are typically preferred to fathers by both sons and daughters.

The sample consisted of 179 male and female college students between ages 17 and 30.

Two scales were used. One was the *PE* (parent evaluation) scale, developed specifically for this study. The other was the *E-F* scale, used in this study as a measure of ideology. The *E-F* scale was administered to each subject three times: as the subject responds for himself; as the subject's mother would respond (phenomenological); as the subject's father would respond (phenomenological). *E-F* difference scores were calculated for each subject: between self and mother, and between self and father.

Results strongly supported the hypothesis. Subjects who gave a parent (mother, father, or both) a relatively high evaluation evidenced relatively close ideological similarity to that parent.

No evidence was found in support of the proposition that opposite-sex parent predilections are responsible for the establishment of parental preference.

Evidence was equivocal relative to the proposition that mothers are typically preferred to fathers.

REFERENCES

1. ACKERMAN, N., & JAHODA, M. Anti-Semitism and Emotional Disorder. New York: Harper, 1950.
2. ADORNO, T., FRENKEL-BRUNSWIK, E., LEVINSON, D. J., & SANFORD, R. N. The Authoritarian Personality. New York: Harper, 1950.
3. ALLPORT, G. The Nature of Prejudice. Cambridge, Mass.: Addison-Wesley, 1954.
4. COOPER, J. B. Mobility anticipation, class assignment, and authoritarianism as field determinants of attitudes. *J. Soc. Psychol.*, 1956, **43**, 139-156.
5. GARDNER, L. P. An analysis of children's attitudes toward fathers. *J. Genet. Psychol.*, 1947, **70**, 3-28.
6. GOUGH, H. G. Studies of social intolerance: I. Some psychological and sociological correlates of Anti-Semitism. *J. Soc. Psychol.*, 1951, **33**, 237-246.
7. HESS, R. D., & HANDEL, G. Patterns of aggression in parents and their children. *J. Genet. Psychol.*, 1956, **89**, 199-212.
8. HOROWITZ, E. L. Development of attitude toward Negroes. *In* G. Swanson, T. Newcomb and E. Hartley (*Eds.*). *Readings in Social Psychology.* New York: Holt, 1952.
9. KRECH, D., & CRUTCHFIELD, R. Theory and Problems of Social Psychology. New York: McGraw-Hill, 1948.
10. NEWCOMB, T. M. Personality and Social Change. New York: Dryden, 1943.
11. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
12. THURSTONE, L. L., & CHAVE, E. J. The Measurement of Attitudes: A Psychological Method and Some Experiments with a Scale for Measuring Attitudes toward the Church. Chicago: Univ. Chicago Press, 1929.
13. WYLIE, P. Generation of Vipers. New York: Farrar & Rinehart, 1942.

Department of Psychology
San Jose State College
San Jose 14, California

The Journal of Genetic Psychology, 1959, **94**, 101-106.

APPARATUS FOR STUDYING VISUAL DISCRIMINATIONS OF MOVEMENT IN SMALL MAMMALS*[1]

Department of Psychology, University of Chicago

LOUIS AARONS AND AUSTIN H. RIESEN[2]

A. INTRODUCTION

Earlier studies of visual movement discrimination in small mammals used apparatus which had limited control of the stimulating conditions (3, 4, 5). The present apparatus was designed for the investigation of movement discrimination in visually deprived cats (1). In exploring the visual ability of these animals it was desirable to vary several parameters of the stimulus. This resulted in a flexible apparatus which is suitable for the following problems: (*a*) moving *versus* stationary figures, (*b*) absolute and relative movement thresholds, (*c*) direction of movement (clockwise *versus* counterclockwise; and expansion *versus* contraction), (*d*) visual intensity, (*e*) flicker, and (*f*) apparent movement. The first three problems may be expanded to include the study of interactions at various levels of luminance. Accurate control of the stimuli are provided for problems (*a*) through (*d*), while (*e*) and (*f*) would require further modifications for more quantitative studies.

B. DISCRIMINATION BOX

The discrimination apparatus is a modified Yerkes-Watson box in which the animal walks into one or the other of the distinguishing cues. Inside dimensions are given in the top outline drawing of Figure 1. The entire structure is made from ½ inch thick plywood and is painted a dull flat black. The inside floor of the box stands 21⅜ inches above the floor of the room. The inside height of the walls of the discrimination chamber is 22 inches, while the height of the walls in the starting box and passageway is 12 inches. The starting box has a hinged plywood cover and a hasp lock, the passageway between the opaque and glass door has a ¼ inch transparent glass top, and the discrimination chamber is not covered. Both the opaque and glass door are operated manually. The partition is made of ¼ inch thick Masonite (12

---

*Received in the Editorial Office on November 25, 1957.

[1] This investigation was supported in part by research grants B-771 and B-772 from the National Institute of Neurological Diseases and Blindness of the National Institutes of Health, Public Health Service.

[2] The authors gratefully acknowledge the technical assistance of Mr. Russel I. Smith and Mr. Friedrich Vojta in the design and construction of the apparatus.

inches length) placed between two ⅛ inch thick pieces of Masonite. The two outer pieces can be slid forward to lengthen the partition. A track for extension is at the top of the partition. The stimulus doors, 8.5 inches x 8.5 inches, are one inch above the floor of the box and 17 inches apart (center to center). They are balanced, hinged, and provided with noiseless locks at their tops. Each door consists of tracing paper sealed between two panes of 1/16 inch transparent glass. The doors serve as screens for projected stimuli.

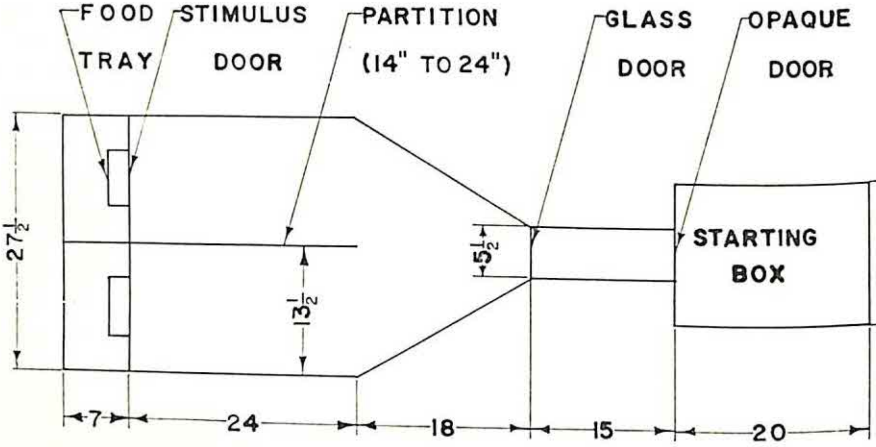

FIGURE 1
Diagram of Two-Choice Discrimination Box

The animal is required to push open the stimulus door, extending its head and shoulders through the door, to obtain a food reward for a correct response.

C. Stimulus Control

The spatial relations of the essential components of the projection system are shown in Figure 2. Two Golde Manumatic projectors (Model No. 300-P-1042), fitted with 9-inch anastigmat coated projection lenses, are mounted behind and above the back end of the discrimination box. The projectors, housed in a dust-free cabinet, are fitted with 2 in. x 2 in. photographic slides of the stimuli. An arrangement of front-surface mirrors stands beneath the projectors. The slides are projected down on the mirrors which reflect the images onto the backs of the stimulus doors.

The reflection system consists of eight mirrors, 8 in. x 10 in., of which there are four for each projector. The arrangement of these mirrors on the turn-table is shown in the insert of Figure 2. The stimulus image is projected onto a center mirror which is fixed at an angle of 45 degrees from the top of the table. This mirror reflects the image sideways to another mirror which is fixed at an angle of 45 degrees to it, is perpendicular to the top of the table, and is situated directly behind the left or right stimulus door. The table rotates on a steel axle which is located at its center of gravity and is supported at each end in an adjustable pillow-block. A spring-tensioned ball-

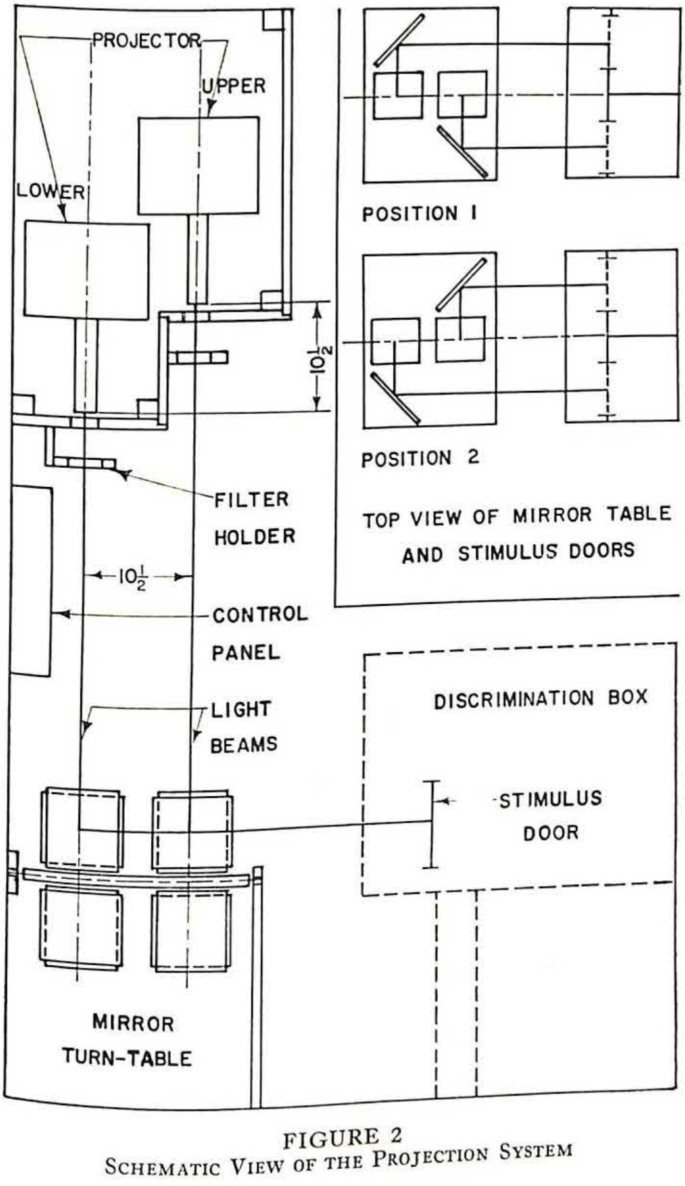

FIGURE 2
SCHEMATIC VIEW OF THE PROJECTION SYSTEM

bearing catch fits a drilled hole in a brass clamp which is located at the side and rear of the table. This catch locks the table in a horizontal plane. The stimuli are switched quickly and quietly from door to door by turning over the table to reverse the reflection system.

Movement of the stimulus image is obtained by constant rotation of the projector. Each projector is fastened with bolts into an aluminum frame. This frame is rotated on the slow shaft of a 2T60 Heller variable speed motor (obtained from Gerald K. Heller, One North Cary Street, Baltimore 23, Maryland) which is equipped with a 600:1 gear reduction unit. This motor can be reversed in its direction of rotation. Rotation of the image about its center is accomplished by the following: (*a*) The motor is securely fastened and supported by the wooden structure enclosing the projectors. (*b*) The center of the motor shaft is joined to the aluminum frame in line with the optical axis of the projector. This axis is made the center of gravity for the mounted projector by adding a counterbalancing weight to the frame. Alignment is maintained by fitting the end of the projector's focussing tube with a plastic guide. This circular guide is firmly attached to the wooden housing. (*c*) Position of the slide is adjusted by use of a modified slide carrier. The projectors are mounted in line with the center of the discrimination box (mid-point of the distance between the stimulus doors). Thus, sounds coming from the rotating motor or projector are not available as cues in the visual discrimination situation. Since it is convenient to use slides bearing identical figures, the stimulus images are equated in size by making the beam length the same for each projector. This is accomplished by mounting one projector higher than the other at a distance equal to the distance between the optical axes of the projectors.

The electrical system of this apparatus is given in the block diagram of Figure 3. Each motor has a separate selenium rectifier power supply which contains a variable voltage transformer (Type 10 Powerstat). Constant voltage transformers (Sola Standard Type 30805) placed in series with each power supply keep the line voltage input within 1 per cent of 115 volts. Rotation speeds are maintained by regulating voltage input to the motor. The reliability of the apparatus for producing a constant speed of rotation is documented in Table I.[3]

Each projector is equipped with a 50 Watt *GE* projector lamp which is operated below its rated voltage. The luminance of the stimulus is controlled by regulating the voltage impressed on the projector lamp. Two variable

---

[3] Calculated values are based on stopwatch readings to the nearest tenth of a second for each revolution of image movement.

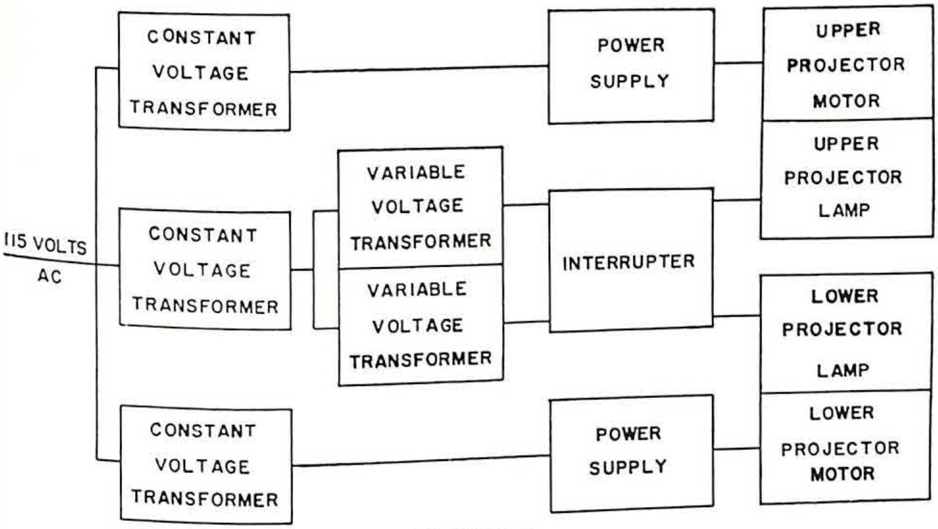
CONSTANT VOLTAGE TRANSFORMER
POWER SUPPLY
UPPER PROJECTOR MOTOR
UPPER PROJECTOR LAMP
VARIABLE VOLTAGE TRANSFORMER
115 VOLTS AC
CONSTANT VOLTAGE TRANSFORMER
INTERRUPTER
VARIABLE VOLTAGE TRANSFORMER
LOWER PROJECTOR LAMP
LOWER PROJECTOR MOTOR
CONSTANT VOLTAGE TRANSFORMER
POWER SUPPLY

FIGURE 3
BLOCK DIAGRAM OF ELECTRICAL SYSTEM

voltage transformers (Type 10 Powerstat) in series with a constant voltage transformer (Sola Standard Type 30806) are used for this purpose. The lamps are wired to this arrangement through commutator rings fastened to the axle upon which the projector rotates. Relatively small changes in luminance are carefully controlled by the use of Wratten neutral density filters sealed in *B* glass or with a sliding wedge. Initial luminance settings are made with a Macbeth Illuminometer. These settings are measured with a

TABLE 1
ROTATION SPEEDS
(Based on Five Measurements)

Power supply dial	Mean of one revolution in seconds	*SD*	Degrees of central angle per second	*RPM*
100	6.26	.055	57.51	9.58
90	8.84	.055	40.72	6.79
80	11.98	.084	30.05	5.01
70	16.25	.055	22.14	3.69
60	20.96	.055	17.18	2.86
50	27.38	.084	13.15	2.19
40	37.20	.122	9.68	1.61
30	56.50	.000	6.37	1.06
20	115.10	1.713	3.13	.52
19	128.58	.487	2.80	.47
18	147.48	.492	2.44	.41
17	168.50	.200	2.14	.36
16	195.66	.192	1.84	.31
15	231.12	3.123	1.69	.26

photoelectric cell connected to a microammeter which is then used for periodic checks.

An interrupter is wired in series with each projector lamp. It contains a synchronous motor with a cam fitted to a microswitch. The cycle of flicker can be varied by changing the size of the notch on the cam. It is important to operate flickering lamps far below their rated voltage.

It has been shown that the monkey sees apparent movement at about the same stimulation parameters as those which are effective with human subjects (2). Our interrupter arrangement can be set up to approximate the optimal interval between successive images in the phenomenon of apparent movement. The phenomenon itself can then be approached by varying the speed of rotation (or actually the distance between successive images).

REFERENCES

1. AARONS, L. Visual movement and intensity discrimination in cats after early restriction and deprivation of pattern vision. Univ. Chicago, Department of Photographic Reproduction, Thesis No. 3532, 1957.
2. DOMINGUEZ, K. E. A study of visual illusion in the monkey. *J. Genet. Psychol.*, 1954, **85**, 105-127.
3. HAWLEY, J. M., & MUNN, N. L. Visual discrimination of movement by white rats. *J. Comp. Psychol.*, 1933, **16**, 137-142.
4. KENNEDY, J. L., & SMITH, K. U. Visual thresholds of real movement in the cat. *J. Genet. Psychol.*, 1935, **46**, 470-476.
5. MORGAN, C. T. The visual discrimination of real movement in the cat. *Psychol. Bull.*, 1937, **34**, 519. (Abstract)

Department of Psychology
University of Chicago
Chicago 37, Illinois

The Journal of Genetic Psychology, 1959, **94**, 107-111.

THE EFFECT OF THE INTRODUCTION OF NOVEL STIMULI UPON MANIPULATION IN RHESUS MONKEYS*

The Radiobiological Laboratory, University of Texas

Richard M. Carr and W. Lynn Brown[1]

A. Introduction

Recent experimentation has attempted to demonstrate that manipulation behavior is dependent upon certain characteristics of the external environment. Harlow, Blazek, and McClearn (2) have postulated that ". . . manipulatory behavior is self-sustaining and is not dependent upon, nor derived from, internal drives such as hunger or thirst, or other incentive systems." Welker (5) gives the following hypotheses concerning the stimulus determinants of play and exploration in chimpanzees:

> (*a*) A relatively new, novel, or different stimulus situation will elicit more exploratory and play behavior than a relatively old, familiar, or unchanged one; (*b*) with repeated presentation, responsiveness to a novel situation decreases not only within any one session, but from one session to another; (*c*) a stimulus object which, when manipulated, causes a change in the situation will elicit more exploration and play than one which does not cause a change when manipulated; (*d*) the animal will preferentially explore and play with those stimulus objects which have greater stimulation value, that is, it will prefer (within limits) larger, brighter, and more heterogenous objects (p. 84).

The present experiment was designed to study the relative effect of novelty of stimulus objects upon manipulatory behavior and the preference for manipulated objects by the rhesus monkey.

B. Method

1. *Subjects*

The subjects were 10 experimentally naïve male rhesus (*Macaca mulatta*) monkeys approximately three years old. The *S*s were fed at 4:00 P. M. daily.

2. *Apparatus*

The *S*s were placed in a metal cage measuring 2 by 3 by 3 feet. A special manipulation box containing three stimulus objects was mounted on the

---

*Received in the Editorial Office on November 27, 1957.

[1] This work was conducted at the Radiobiological Laboratory of The University of Texas and the United States Air Force, Austin, Texas, and was supported (in part) with funds provided under Contract AF 41(657)-149 and the School of Aviation Medicine.

barred front of the cage. The manipulation box (28 by 12 by 4 inches) was constructed of ⅜-inch plywood. Three ¼-inch holes were drilled through the front of the box, one in the center and one 7 inches from each end. Springs were inserted in these holes and were attached to an automatic recording device within the box and to the stimulus object at the other end. A pressure of 50 grams upon any of the three stimulus objects would activate the recording device. The maximum weight of any one of the stimulus objects was 25 grams.

3. *Procedure*

The *S*s were tested for six 5-day periods; the duration of each daily session was 10 minutes. For each group of 5-day periods, three stimulus objects constructed of the same material and differing in shape and color were employed. Records of the number of manipulations and the amount of time spent manipulating at each session were kept for each animal. All *S*s were run in close succession between 8:30 and 11:00 A. M. daily with the order of testing the *S*s being randomized.

C. RESULTS

The statistical significance of all results reported below were tested by Kendall Coefficient of Concordance: *W* (4).

Figure 1 shows the total number of manipulations for all 10 *S*s over each day of the 5-day period with respect to each type of manipulanda. With the exception of metal and cork, the curves demonstrate a maximum of manipulation on the first day of each 5-day period with a sharp decrease on the succeeding days, reaching a minimum of the fifth day. The relative heights of each of the curves demonstrates a significant difference in the *S*s preference for the different objects ($p < .05$). Object preference is defined as the total amount of manipulation of the stimulus objects. The order of preference of the *S*s for the manipulanda is wood, rubber, cork, wire, plastic, and metal. This is comparable to the McDowell results (3).

Figure 2 represents the total amount of time spent manipulating each of the stimulus objects by all 10 *S*s for each day of testing. Again with the exception of cork, there is a close resemblance between Figures 1 and 2. In this case, the difference in the amount of time spent manipulating each of the different stimulus objects is significant at the .01 level of confidence.

The last point on Figures 1 and 2 represents a 1-day retest of all *S*s on wood. This retest was deemed necessary since the over-all trend of the first day trials of both figures over the total 30-day period suggested that the manipulatory behavior reaches a maximum and then gradually declines. The

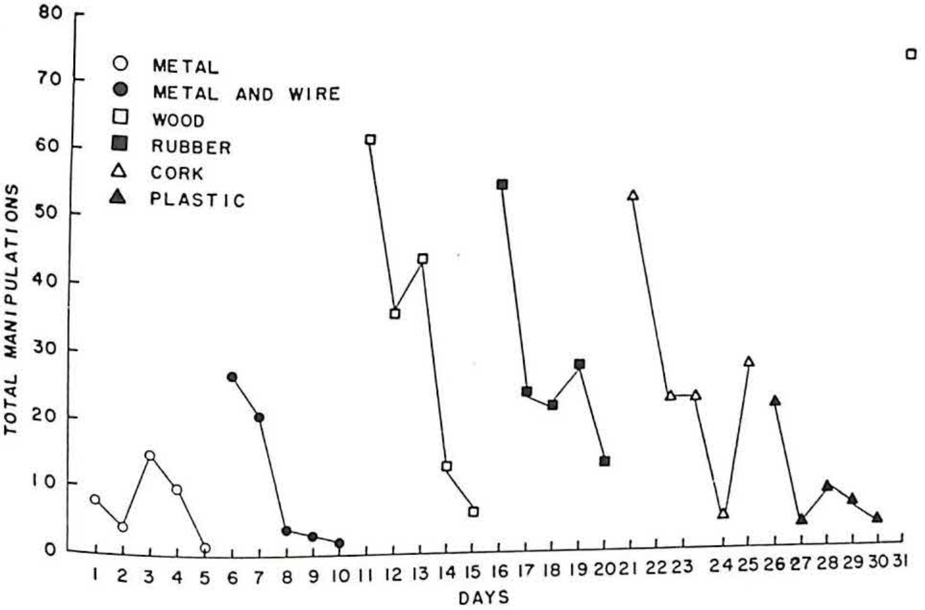

FIGURE 1
Manipulation of Stimulus Objects by All Subjects during the 31-Day Testing Session

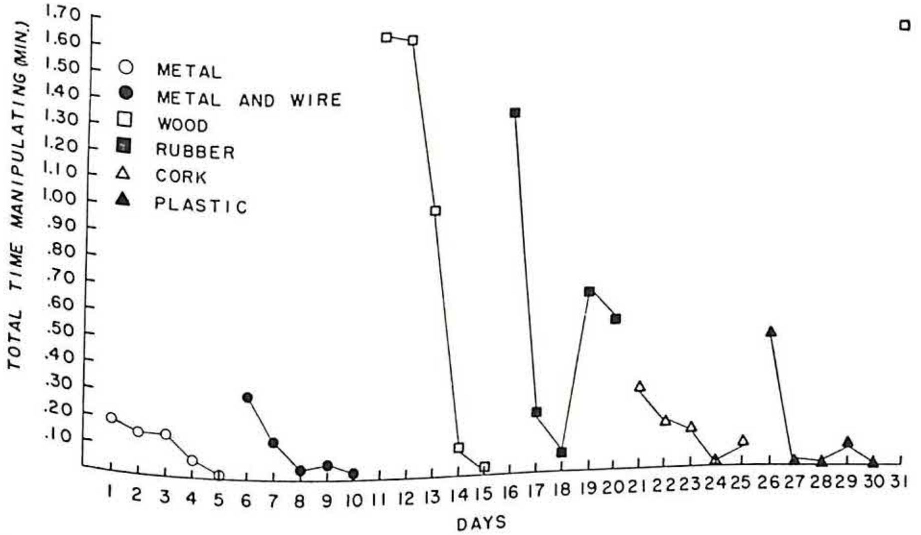

FIGURE 2
Total Time Spent Manipulating Stimulus Objects by All Subjects during the 31-Day Testing Session

height of this retest day suggests that the over-all curve for the 30-day period was a function of manipulatory behavior and not the order of presentation.

D. Discussion

The results of this experiment demonstrate that the four hypotheses suggested by Welker (4) to explain the exploratory and play behavior of chimpanzees also pertains to the manipulatory behavior of the rhesus monkey. The results of this experiment suggest the following hypothesis: The greater the amount of physical change the monkey is able to produce in the stimulus object, the greater the amount of manipulation, the greater the amount of time spent manipulating the stimulus object, and the slower the extinction.

A maximum of manipulation occurs on the first presentation of each of the stimulus objects. The next presentation of the novel stimulus object again affects the receptors, and there is a second session of manipulation, but this time there is less manipulation than on the preceding presentation. If the *S*s are unmotivated, Berlyne (1) would predict such manipulation will recur on succeeding presentations until manipulations of the stimulus object cease altogether.

The variability occurring in the fourth and fifth days has led the authors to believe that the preference for stimulus objects determines the duration of the novelty. Harlow, Blazek, and McClearn (2) have assumed that manipulatory behavior is self-sustaining. It appears to us it is self-sustaining only because the novelty which results from manipulation is rewarding. The novelty of the situation attracts the *S* to respond to the manipulanda, and once this is begun, the ensuing behavior is self-sustaining. Experiments of this type should be carried over long periods of time using each of the stimulus objects.

E. Summary

Ten experimentally naïve male rhesus (*Macaca mulatta*) monkeys were tested on a special manipulation box for 10 minutes each day for five days on each of six different stimulus objects. The objects were made of metal, wire, wood, rubber, cork, and plastic.

The results demonstrate that manipulatory behavior is directed by the novelty of the situation with the greatest manipulation occurring with those stimulus objects which the *S* can physically alter most in appearance. The order of preference for both the amount of manipulation and the amount of time spent manipulating was wood, rubber, cork, wire, plastic, and metal. It appeared that manipulatory behavior in the present experiment was self-sustaining and the novelty resulting from manipulation is rewarding.

REFERENCES

1. BERLYNE, P. E. Novelty and curiosity as determiners of exploratory behavior. *Brit. J. Med. Psychol.*, 1950, **41**, 68-80.
2. HARLOW, H. F., BLAZEK, NANCY C., & McCLEARN, G. E. Manipulatory motivation in the infant rhesus monkey. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 444-448.
3. McDOWELL, A. A. Normative behavior of the caged macaque (*Macaca mulatta*). (Unpublished Master's Thesis, University of South Dakota, 1953.)
4. SIEGAL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
5. WELKER, W. I. Some determinants of play and exploration in chimpanzees. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 84-89.

The Radiobiological Laboratory
University of Texas
Austin, Texas

The Journal of Genetic Psychology, 1959, **94**, 113-130.

THE MIND AS AN OPTICAL SYSTEM: A STUDY OF PERCEPTUAL REPRODUCTION IN THE NORMAL AND MENTALLY RETARDED*

St. Christopher's Hospital for Children, Philadelphia

G. H. Zuk

A. Introduction

The intent of this study was to elaborate on the mental process which intervenes between the input of a visual stimulus and its reproduction on paper. Groups of retarded and nonretarded individuals were compared in their respective capacities to reproduce such visual stimuli as were shown them.

Freud (2), in some of his early work in personality theory, supplied a model encompassing the structure of the mind. In his analysis of dreams he described an apparatus with the physical characteristics of an optical system, such as a telescope or microscope. He was led to conclude that the mind behaved in many respects as if it were a reflex apparatus—a microscope, for example—with a perceptual lens at one end and a motor lens at the other. Between the lenses, according to Freud, there existed additional lens systems which could alter the course of incoming signals. He believed that in sleep the motor end of the apparatus was sealed off. Stimuli originating at the perceptual end were not allowed free access; they were shunted back toward the perceptual end. He proposed that the dream was actually the product of this regressive process. As the stimuli were returned toward the perceptual end of the apparatus they passed through systems which evoked the images which constituted the dream. In seldom quoted passages, Freud described this early model:

> . . . we shall do no more than accept the invitation to think of the instrument which serves the psychic activities much as we think of a compound microscope, a photographic camera, or other apparatus. The psychic locality, then, corresponds to a place within such an apparatus in which one of the preliminary phases comes into existence. As is well known, there are in the microscope and the telescope such ideal localities and planes, in which no tangible portion of the apparatus is located. . . .
>
> Accordingly, we conceive the psychic apparatus as a compound instrument, the component parts of which we shall call instances, or, for

---

*Received in the Editorial Office on November 29, 1957.

> the sake of clearness, systems. We shall then anticipate that these systems may perhaps maintain a partial orientation to one another, very much as do the different and successive systems of lenses of a telescope. Strictly speaking, there is no need to assume an actual spatial arrangement of the psychic system. It will be enough for our purpose if a definite sequence is established, so that in certain psychic events the systems will be traversed by the excitation in a definite temporal order. . . .
>
> The first thing that strikes us is the fact that the apparatus composed of the psychic systems has a direction. All our psychic activities proceed from inner or outer stimuli and terminate in innervations. We thus ascribe to the apparatus a sensory and a motor end; at the sensory end we find a system which receives this perception, and at the motor end another which opens the sluices of motility. The psychic process generally runs from the perceptive end to the motor end. . . . (2, pp. 392-393.)

The reader is invited to refer to the diagrams shown by Freud (see 2, pp. 394-396) in order to picture the model that has been described in the present context.

What direction did the perceptual stimuli take as they entered the psychic apparatus? Freud concluded that when such stimuli entered they left definite traces of themselves in systems which lay just behind the perceptual lens; he called these "mem" (for memory) systems. From the "mem" systems, the stimuli poured forward into other systems, which were described by Freud as follows:

> The last of the systems at the motor end we call the preconscious (Pcs.) to denote that the exciting processes in this system can reach consciousness without any further detention, provided certain other conditions are fulfilled, e.g., the attainment of a definite degree of intensity, a certain apportionment of that function which we call attention, etc. This is at the same time the system which holds the keys of voluntary motility. The system behind it we call the unconscious (Ucs.), because it has no access to consciousness except through the preconscious, in the passage through which the excitation must submit to certain changes (2, p. 397).

The purpose of this description of the Freudian model has been to establish a theoretical framework within which the findings of the study to be outlined can operate. In the writer's opinion, the seldom used Freudian model of the mind as possessing optical properties brings an order to the present data that no other model of his acquaintance could provide.

The design of this study was developed in accord with the general research procedure outlined by the writer (4) in a recent paper. In it he described a variant of the traditional longitudinal and cross-sectional approaches, called "condition-analysis," in which standard material is readministered to subjects in a series of conditions slightly varying from each other along some dimension—intensity, for example. This approach permits a high degree of experimental control to be maintained and yet allows sufficient flexibility to reduce the level of artificiality in the experimental procedure.

B. DESCRIPTION

1. *Subjects*

Three groups of subjects were utilized in the present study:

1. Approximately 125 institutionalized patients in a hospital for the mentally retarded.[1] A wide variety of diagnostic categories of endogenous and exogenous disorders were included in the group. The so-called "undifferentiated" group was represented in the greatest number, inasmuch as this group tends to occupy the highest *IQ* range in a population of mentally retarded. These individuals are called "undifferentiated" for the reason that a physical basis for their retardation is not known. Very few Mongols were present in the group, inasmuch as they tend to occupy an *IQ* range too low for the purpose of the present study.

About 80 per cent of the group involved in the testing procedure were adults, that is, were over 17 years of age. Approximately 20 per cent were under 17 years of age, but none of this group was younger than 13 years. Males and females were represented in the total group in the ratio of about 2 to 1.

Mental ages for the retarded group occupied the range from 6-0 to 12-0, the bulk however within the 8-0 to 9-0 class. This meant that *IQ*'s for the group ranged from about 50 to 70. A subject with less than an *MA* (mental age) of 6-0 could usually not perform the drawing task required.

2. About 45 adult psychiatric technician trainees at Pacific State Hospital who were students in a psychology course being instructed by the present writer. Women predominated in the group, the ratio being about 4 to 1. Ages varied from about 18 to 55 years. Academically, these individuals were required to have completed a high-school education. A few members were attending a nearby junior college. *IQ*'s were not available to the writer, but on the basis of his knowledge of the group's performance on different tests he administered he would estimate the range as being between 85 and 130.

3. Fifteen graduate students participating in a seminar in mental retardation conducted by the writer at a local state college constituted the

---

[1] The writer wishes to thank the following individuals at Pacific State Hospital in Pomona, California, for their generous help in obtaining the necessary data: Dr. George Tarjan, medical director, Dr. Anna Shotwell, chief psychologist and Mrs. Mayme Dinsmore, principal of the School.

third group of subjects. These individuals were all adults, about evenly divided as to sex, and had worked or were working in the field of special education. Most had received their Master's degree; the remainder were about to receive it.

2. *The Test Material*

For the purpose of this study, the writer constructed a set of 10 geometrical figures (see Figure 1). The rationale for the particular series developed was simply to include figures which would require the subject to reproduce straight lines, wavy lines, circles, dots, different-sized angles, triangles, and so on. In other words, the object was to make a set which required subjects to reproduce a wide variety of geometrical shapes.

Each of the figures was drawn in India ink on a 5″ x 8″ white card. Drawings were made of sufficient size and clarity to be seen easily at any point within the halls and classrooms used in the testing procedure. Any request for greater proximity on account of impaired vision was of course immediately satisfied.

The figures of the set varied in length and width. Of the four figures intensively measured in the present study (see 1, 4, 7, and 10 in Figure 1), 10 was the longest (4⅜″ by 2″), 4 was the shortest (2¼″ by 1¼″), 7 was next longest (3⅞″ by 1¾″), and 1 was next shortest (3½″ by 1⅝″).

3. *The Variable and Unit of Measurement*

In order to hold the subjective element to a minimum, the writer determined to measure only selected physical dimensions of figures reproduced. Length and width of the figures appeared to be the dimensions most easily accessible to measurement. To have taken measurements on both of these would, however, have proved unduly burdensome in view of the many thousands of readings that would have been required. The writer selected the dimension of length for the reason that it seemed to him (after examining several test protocols) that a more equitable dispersal of scores would be obtained. In other words, it looked as if measurement along the dimension of length would produce a spread of scores of sufficient size and with enough central tendency to permit statistical handling of the raw scores directly. It was also a reasonable assumption that there would be at least a moderate positive correlation between measures of the length and width dimensions; that is, that the longer a particular figure, the wider also.

Length of the figures was measured by ordinary ruler. A system of scaling was constructed in which ⅛″ represented 1 unit. For example, a figure 1″

in length was given a score of 8; a figure drawn 2″ in length, a score of 16; a figure drawn 3″ in length, a score of 24, and so on. Individual figures were measured from the point drawn closest to the left hand side of a sheet of 8½″ by 11″ yellow paper to the point drawn closest to the right hand side.

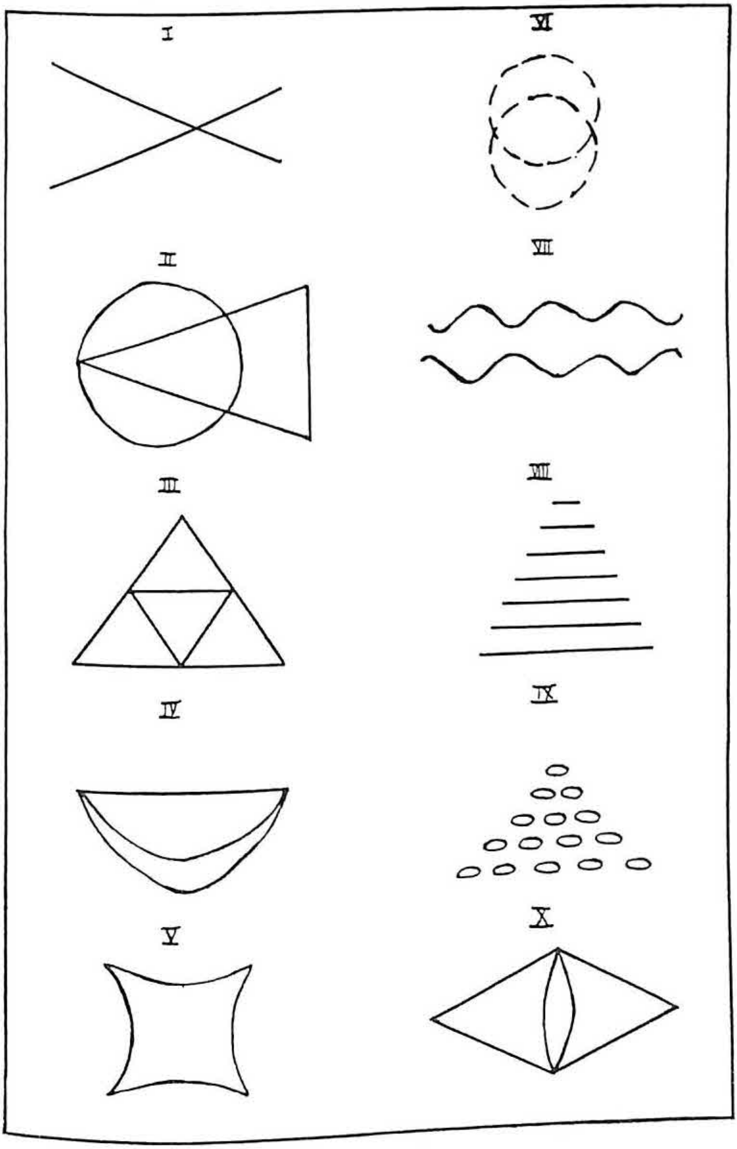

FIGURE 1
FIGURES 1-10

Because even length measurements of the series of 10 figures would have required over 6,000 readings, it was decided by the writer to select what might on rational grounds be called a "representative sample" of the series. Inasmuch as they occupied equidistant points in the series, Figures 1, 4, 7, and 10 were selected to be measured.

4. *Measurers*

Figures reproduced by the mentally retarded group were measured in terms of their length by the 15 graduate students who also took part in this study as subjects. The 125 test protocols derived from the mentally retarded group were distributed equally among the graduate student group. Inasmuch as the graduate group had *previously* reproduced the series of figures in the study as subjects, their test protocols were not in any way influenced by their knowledge of the dimension along which measurement was to occur.

Length measurements of the test protocols of the psychiatric technician-graduate student combined group of about 60 subjects were done by the present writer.

5. *Conditions of Reproduction*

The figures were reproduced by the subjects under the following conditions:

1. (Mentally retarded group). In groups of between 10 and 30, subjects were passed 4 sheets of yellow paper, 8½″ by 11″. They had previously been informed by the regular class teacher that they were going to be asked to "draw" for the present writer.

The writer explained to the subjects (in a typical school classroom) that they were to copy each of the figures that he would show them on a single sheet of the yellow paper. He held the set of cards in his hands to give the group an idea of how many were to be copied. The writer then said that he would show them each card. He asked the group to look closely at the card. They were told that after a short time the card would be withdrawn and that they would then have as much time as they needed to draw the card they had seen. They were informed that they had only a single sheet of paper on which to draw all of the cards. Under these directions the presentation and reproduction of the series of figures proceeded, *each figure being viewed for about 10″*. Papers were collected at the end of the series.

The subjects were asked whether they thought they "could do better." Many, of course, replied that they could. They were then readministered the series of figures (to be drawn on a second sheet of paper). In this instance, *each figure was exposed for a period of about 5″*. Papers were collected.

Subjects were asked once again whether they thought they "could do better." Although by this time some were growing restless with the task, a few answered in the affirmative. The series was readministered, drawn on the third sheet of paper supplied the group. *Figures were exposed in this condition for approximately 2″*. Papers were collected.

Subjects were then asked what they thought the fourth sheet of paper was for. In most cases, at least one person "guessed" that the group would be asked to reproduce the figures from memory. In this condition, subjects were informed before they began that they would have as much time as they wanted with the task and that only when they thought that they could not recall any more figures should they hand in their papers. Conditions were contiguous in time; all were completed in about 30 minutes.

The reader, at this point, may wish to inquire: was copying from each other extensive? The writer would answer that while some copying of figures was observed it had little or no affect on the results of this study, inasmuch as *length* (and not shape) was the dimension being tapped. The reader might also ask: were there many subjects with visual defect present in the group? The writer would reply that known visual defect was no more prevalent in the retarded than in the nonretarded group. It should be re-emphasized that it was not necessary for the subject to reproduce a perfectly shaped figure since shape was not a variable under consideration here. In view of the fact that the subjects were free to examine the figures which they reproduced at any distance from the tip of their noses to full arms length, it is assumed that only the grossest of visual defect would affect consistently the length of figures reproduced.

2. (Psychiatric technician group). The same general procedure was followed in this group as in the mentally retarded group. Exceptions in procedure are noted as follows: because of time restrictions, only the 10″, 2″ and memory conditions were included in the total procedure—the 5″ condition was omitted; and there was a period of about 15 minutes elapsing between the 2″ and memory conditions because of a break-period in the class. Therefore, the period between the 2″ and memory conditions for the technicians was about 5 times longer than it was for the retarded group. Time of test procedure: about 45 minutes.

3. (Graduate student group). The procedure was almost identical with that described for the mentally retarded group. Conditions were contiguous in time. Conditions were set for the 10″, 5″, 2″ exposure and from memory. Time of test procedure: about 30 minutes.

6. *Summary of Design*

Three groups, two of normal and one of retarded intelligence, were requested to reproduce a set of 10 figures which was exposed to them in orderly succession for intervals of 10″, 5″, 2″ and from memory. Thus four separate reproductions of the test series were obtained from which comparisons could be made.

Four of the figures of the set were subjected to intensive measurement along the dimension of length. It was considered that the subjective element in measurement of length would be held to a minimum.

The design permitted "within-characteristics" of the mentally retarded and nonretarded groups to be elicited and also "between-characteristics."

C. RESULTS

Table 1 lists *N*'s, means, standard deviations, mean differences, and *t*-tests (with their levels of significance) for length measurements of four of the set of 10 figures reproduced by mentally retarded subjects (approximately 125) and two groups of nonretarded individuals (appproximately 60 in the combined psychiatric technician-graduate student group). As mentioned previously, the psychiatric technicians were not exposed to the 5″ condition and thus such entries as would be made in the table had they been so exposed are not listed. The 10″ condition was used as the standard for comparison with results obtained in other conditions, i.e., results obtained from the 10″ condition were compared with results obtained from the 5″, 2″ and memory conditions by means of *t*-tests to test the significance of mean variation.

Centering attention initially on material presented in Table 1 relevant to the mentally retarded group, the reader will observe that for Figures 1 and 4 the mean length tends to decrease in the successive conditions, the clearest drop occurring between the memory condition as compared with the three others. The mean difference between lengths of Figure 1 in the 10" and memory condition is significant at less than the .01 level. On Figure 10 there appears to be no tendency for decrease in length in the first three conditions, although a statistically significant drop occurs in the memory condition (*P* less than .05). Figure 7 appears to reverse the direction established by measurement of the three other figures, its length *increasing* in the first three conditions. However, the trend established with respect to the memory condition is continued, inasmuch as the mean length for Figure 7 is that condition declines below levels found in three previous conditions.

These results indicate that (for the mentally retarded group) decreasing the period of exposure of the cards from 10″ to 5″ to 2″ had little or no

TABLE 1

MEANS, STANDARD DEVIATIONS, N'S, MEAN DIFFERENCES, AND THEIR LEVELS OF SIGNIFICANCE FOR THE MENTALLY RETARDED AND NORMAL GROUPS WITH RESPECT TO THE VARIABLE LENGTH OF FIGURES REPRODUCED IN SUCCESSIVE CONDITIONS

Mentally retarded group

	Figure 1**				Figure 4				Figure 7				Figure 10			
Statistics	10"	5"	2"	Mem	10"	5"	2"	Mem	10"	5"	2"	Mem	10"	5"	2"	Mem
N	126	126	126	121	125	126	124	92	126	126	126	103	119	120	122	102
M	9.83*	9.75	9.66	8.02	8.78	8.56	8.27	8.18	13.11	13.92	13.88	12.95	13.41	13.42	13.40	12.02
S.D.	6.73	5.46	6.49	5.47	4.33	3.89	4.54	5.15	6.63	6.00	6.83	6.48	7.09	6.49	7.06	6.77

Mean differences for mentally retarded group

Conditions	Md.	t.	P.	Md.	t.	P.	Md.	t.	P.	Md.	t.	P.
10"–5"	.08	.16	—	.22	.61	—	.81	1.56	—	.01	.02	—
10"–2"	.17	.29	—	.51	1.19	—	.77	1.38	—	.01	.02	—
10"–Mem	1.81	3.07	<.01	.60	1.05	—	.16	.07	—	1.39	2.07	<.05

The normal group

Statistics	10"	2"	Mem	10"	2"	Mem	10"	2"	Mem	10"	2"	Mem
N	60	60	60	60	60	57	60	60	59	59	60	57
M	17.15	14.50	14.33	12.45	10.43	10.95	20.22	20.17	20.28	17.66	17.98	16.72
S.D.	9.15	5.80	5.66	4.66	3.11	3.73	5.82	6.25	7.13	5.95	6.49	5.83

Mean differences for normal group

Conditions	Md.	t.	P.	Md.	t.	P.	Md.	t.	P.	Md.	t.	P.
10"–2"	2.65	2.82	<.01	2.02	4.04	<.01	.05	.06	—	.32	.41	—
10"–Mem	2.82	2.94	<.01	1.50	2.78	<.01	.06	.07	—	.94	1.31	—

*Indicates that average length of this figure reproduced by mentally retarded group in 10" condition was ⅛" by 9.83 or almost 1¼".

**Actual length of figure 1 on 8" by 5" card was 3½" (or 28 units); of figure 4, 2¼" (or 18 units); of figure 7, 3⅞" (or 31 units); of figure 10, 4⅜" (or 35 units).

effect upon the length of the figures reproduced; but that reproducing figures from memory *definitely tended to reduce the length of all figures measured.* From trends apparent (e.g., the length of Figure 1 slightly decreasing in the first three conditions, the length of Figure 4 slightly decreasing, the length of Figure 7 increasing slightly, and the length of Figure 10 remaining about constant) the writer concludes that a certain positional effect was operating in addition to an effect having to do with special characteristics of the geometric configurations themselves. This effect appears to be relatively minor, however, and the most pertinent finding that emerges is that memory affects the length dimension in the perceptual reproduction of mentally retarded subjects, causing shorter figures to be reproduced.

In the combined nonretarded group (psychiatric technicians and graduate students) memory, surprisingly enough, appeared to exert no such diminishing effect as in the case of the mentally retarded group.

For the nonretarded group on Figure 1 there is a significant mean decrease in lengths produced in the 10″ and memory conditions (P less than .01) but there is also a significant decrease in lengths produced in the 10″ and 2″ conditions, which suggests that in this instance the reduction in the length of the figures was due more to a learned adjustment than to the memory effect. The same general situation seemed to hold with respect to Figure 4. By Figure 7, however, the learned adjustment seems to have become established and there is little or no variation in lengths in the successive conditions. This situation holds also with respect to Figure 10. By "learned adjustment" is really meant a correction for the positional effect which was posited as affecting to a small degree the lengths of the figures as they were drawn in sequence and also as affecting length between successive sets.

Thus memory, in the nonretarded group, did not influence the length of figures reproduced. Once the nonretarded group had made a learned adjustment to the positional effect imposed by administering the test instrument in series, no significant fluctuations in length occurred. In addition the results suggest that certain characteristics of the geometric figures themselves seemed to have had less effect in altering the length of figures in any particular direction in the successive conditions than was the case with the retarded group.

In Table 2 there emerges the fact that the normal group produced, regardless of the particular condition, significantly *larger* figures than the retarded group. In no comparison made was the mean difference less statistically significant than at beyond the .01 level. The length of the figures

TABLE 2

COMPARISON OF THE MENTALLY RETARDED AND NORMAL GROUPS WITH RESPECT TO LENGTH OF FIGURES REPRODUCED IN SUCCESSIVE CONDITIONS

Normal vs. retarded group	Figure 1			Figure 4			Figure 7			Figure 10		
Conditions	Md.	C.R.	P.	Md.	C.R.	P.	Md.	C.R.	P.	Md.	C.R.	P.
10"	7.32	5.50	<.01	3.67	5.10	<.01	7.11	7.41	<.01	4.25	4.21	<.01
2"	4.84	5.09	<.01	2.16	3.79	<.01	6.29	6.23	<.01	4.58	4.24	<.01
Memory	6.31	7.09	<.01	2.77	3.85	<.01	7.22	6.39	<.01	4.70	4.61	<.01

reproduced by the normal group were in every case from about ¼ to ½ larger than those reproduced by the retarded group.

Table 3 summarizes the data by means of Pearson product-moment correlations between lengths of identical figures reproduced in successive conditions. The correlations give an estimate of the level of consistency or stability of the figures with respect to the length variable. Correlations for the retarded group ranged from .35 to .66. Generally speaking, correlations obtained between the 10″ and 5″ conditions were somewhat stronger than those obtained between the 10″ and 2″ or between the 10″ and memory conditions.

TABLE 3
CORRELATIONS BETWEEN LENGTHS OF IDENTICAL FIGURES REPRODUCED BY THE MENTALLY RETARDED AND NORMAL GROUPS IN THE SUCCESSIVE CONDITIONS

	Figure 1		Figure 4		Figure 7		Figure 10	
Conditions	Retarded	Normal	Retarded	Normal	Retarded	Normal	Retarded	Normal
10″–5″	.56	—	.53	—	.57	—	.66	—
10″–2″	.55	.60	.42	.67	.56	.51	.57	.54
10″–Mem.	.55	.59	.35	.55	.59	.50	.55	.59

Correlations of figures in successive conditions for the combined nonretarded group ranged from .50 to .67. In other words, both the retarded and nonretarded groups maintained roughly similar levels of consistency in reproducing identical figures in the different conditions.

Table 4 lists Pearson product-moment correlations between different figures reproduced in the memory condition only. This procedure provided some check on the statistical independence of the four figures with respect to the length variable. For the mentally retarded group, inter-correlations ranged from —.12 to .26. For the nonretarded group inter-correlations ranged from .38 to .65. Seemingly, the similarity between lengths of the four figures was somewhat greater for the nonretarded than for the retarded group.

TABLE 4
INTERCORRELATIONS BETWEEN FOUR OF THE FIGURES REPRODUCED BY THE MENTALLY RETARDED AND NORMAL GROUPS IN THE MEMORY CONDITION

	Figures					
Groups	1–4	1–7	1–10	4–7	4–10	7–10
Mentally retarded	.18	—.12	.23	—.08	.26	.03
Normals	.60	.48	.65	.57	.62	.38

It should surprise few that the nonretarded group recalled a significantly greater number of figures in the memory condition than the retarded. The mean number recalled by the nonretarded was 9.59, with a standard deviation of .67. The mean number recalled by the retarded group was 8.05, with a standard deviation of 1.60. The mean difference was 1.54 which, when divided by the appropriate error term, gave a critical ratio of 9.06 (significant at beyond the .01 level). The nonretarded group might have achieved an even higher mean level had the psychiatric technicians been administered the memory condition immediately following the 2″ condition, as was the case with the retarded group and the graduate students. But, as reported earlier, a period of about 15 minutes elapsed between the 2″ and memory conditions.

The unhappy state of diagnosis in the area of mental deficiency in addition to the fact that some diagnostic groups were represented in the total retarded group in only very limited numbers made it practically impossible to obtain definitive results with respect to the performance of different subgroups of retarded on the test series. One grouping made compared the reproductions of the retarded having the lowest mental ages (*N* of about 17, *MA*'s between 6-0 and 7-0) with the reproductions of the graduate student group (*N* of 15, the group that possessed the highest level of intellectual performance). This comparison tended to bear out even more strongly the type of relationship that was found when the retarded group as a whole was compared with the combined normal group, i.e., that there was a positive relationship between level of intelligence and length of figure reproduced. In a subgrouping of the retarded in terms of diagnostic categories, the writer compared 14 subjects judged by examining psychologists as either severely schizoid or schizophrenic with 10 subjects who showed a history of epileptic seizures. Because of the limited number of subjects available, the results of this comparison were inconclusive and shall not be reported in this context. Obviously, this would be the type of procedure to be followed by individuals interested in determining whether the test series would be of value in differentiating mentally retarded individuals by diagnostic category. Much larger groups of subjects would be required in order for definitive results to be obtained. All that can be said at this time is that, in the writer's opinion, the test series described here would appear to be of value in differential diagnosis in the area of mental retardation. It should be emphasized that the writer does not so much recommend a test series per se as he does a test series applied within a definite methodological framework, such a framework as he has called "condition-analysis."

D. Discussion

1. *Concept of the Mind as an Optical Instrument*

As shown in Freud's model of the psychic apparatus (2, p. 396), percepts entering the perceptual (*P*) end are transmitted directly to "mem" systems which make a trace. According to the model, percepts leaving the "mem" systems then enter the area of the unconscious (*Ucs.*), proceed to the preconscious (*Pcs.*) and finally achieve outward expression through the motor (*M*) system. It was Freud's idea that any of these systems could affect the course and character of the percept, especially so the preconscious which operated as the strongest censor of percepts. He expressed the idea that each system of the psychic apparatus possessed optical properties; in other words, that each system had the capacity to affect incoming stimuli in precisely the same manner that an optical instrument (such as a microscope) could affect incoming light rays. Let us consider this model in relation to the findings of the present study.

We may assume that the percepts involved in this study (i.e., the geometrical figures) traversed their course within the psychic systems of the subjects, from the *P*-end to the *M*-end. Percepts were converted into their motor equivalents at the *M*-end by motor reproduction on paper.

That the nonretarded should reproduce significantly longer percepts than the retarded now becomes especially noteworthy, for this finding compels the interpretation that such differences are related to differences within the psychic structures of the groups. In other words, one might deduce that special distortions occurred in the transmission of percepts in the psychic systems of retarded individuals which did not occur or occurred in another way in nonretarded individuals.

Which systems were responsible for producing the difference? Certainly not the "mem" systems alone, since it was shown that the reproductions of the retarded were significantly smaller than the nonretarded even when the memory function was hardly operating (as in the 10″, 5″ and 2″ conditions).

We look instead to disturbances in other psychic systems to account for this result, more likely at the *P*-end of the apparatus. On the other hand, some impairment in the "mem" systems is actually reflected in the finding that retarded subjects consistently reproduced smaller figures in the memory condition than they had in the three previous "non-memory" conditions.

The nonretarded group tended to reproduce more accurately in the sense that the *length of their figures more closely approximated the actual length of the figures exposed to them on cards.* They seemed to be less influenced by positional effects and by special gestalt characteristics of the figures themselves. In terms of the Freudian model, the percepts seemed to pass through the various psychic systems in a less distorted fashion.

Two functional characteristics differentiating the psychic systems of the retarded and nonretarded seem to emerge from the findings: (*a*) that the various systems are in a sense "weaker" in the retarded, i.e., that the percept loses greater amounts of energy as it passes through, and (*b*) that the various systems contain more structural flaws in the retarded which seem further to distort the percept. It was noted, for example, that only in retarded individuals with mental ages less than 6-0 were gross distortions present; e.g., Figure 1 reproduced as a "cross," Figure 4 as a "bowl," Figure 10 an "eye." Bender (1) has commented on this predisposition of the retarded to elaborate on and grossly misrepresent percepts.

In the light of these findings, the Freudian model appears to make possible also a structural-dynamic explanation for the concrete-mindedness of the brain-damaged and mentally retarded, a phenomenon which has been so excellently described by Goldstein. Let us assume that in individuals with mental ages below 6 years, percepts are grossly weakened and distorted as they are received at the *P*-end of the psychic system. The mem-systems are, as a result, impressed with a weakened signal. Suppose then that the system immediately adjacent to the mem-systems (the unconscious) places certain requirements on incoming signals, one being that they possess intelligibility (that is, they must "make sense") before they are accepted and carried forward into the preconscious. And suppose that in order to make the percept intelligible to the unconscious the mem-systems must "tag" or label (that is, concretize, in Goldstein's sense) the weakened signal by investing it with meaning as an object familiar in experience (such as an "eye" for Figure 10, a "bowl" for Figure 4). This investment of meaning then would serve to re-energize the weakened percept and insure its transmission.

What Goldstein has termed concreteness in describing the thinking of brain-damaged individuals may be accounted for in a structural-dynamic conceptual scheme. In this respect, F. C. Bartlett's classical study on remembering supplies some light. Using normal subjects he found stories and pictures alien to their culture reproduced in a manner which modified the content in the direction of the familiar. We would explain this finding in much the same way that we have explained parallel findings described in this study: that mem-systems, striving to transmit an intelligible percept to other systems within the total psychic system, invest an ambiguous signal which has been weakened at the *P*-end with concrete significance, thus re-energizing the signal and insuring its transmission through the other systems.

It is the belief of the writer that a model of the psychic apparatus as possessing properties similar to those of optical systems is most helpful in explaining some of the disturbances which are commonly seen in the mentally retarded as well as in the mentally disturbed. It is an early Freudian model which placed great emphasis upon the structural integrity of the various systems composing the psychic apparatus, in contrast to the later Freudian thinking which tended to stress more strongly the dynamic content of the various systems. In other words, as this writer sees it, the earlier model called for a study of the constitutional weaknesses and strengths of the various psychic systems; the later model (or models, really) called for a study of the weaknesses and strengths *of various ideas or thoughts* resident in different systems. As an example of thinking consistent with the later model, the following: a very powerful wish-impulse (idea or thought) present in the id is said to struggle for expression and finally to achieve it by overcoming ego ideas. As an example of thinking consistent with the earlier model, the following: structural defects in the unconscious (*Ucs.*) and in the preconscious (*Pcs.*) systems permit the passage of wish-impulses which would otherwise never have been allowed to move freely from one system to another.

This difference in models is more than just a semantic one. It is a difference which has resulted in a possible over-emphasis on thought-contents of the psychic systems; it has resulted in a neglect of structural characteristics.

2. *The Condition-Gestalt in Clinical Use*

The test series described in this paper, which for convenience may be labelled the Condition-Gestalt, would appear to have definite clinical-diagnostic value. It was demonstrated that it possessed differentiating power with respect to the variable of intelligence. Only the surface has been scratched, however, in terms of establishing the Condition-Gestalt as a component part of a general psychological battery. Further work is necessary in order to bring this to pass.

One question which the interested investigator might wish to pursue in using the Condition-Gestalt in experimentation into its clinical value is: what is the significance of "small" figures (i.e., figures the size of those characteristically reproduced by retarded individuals) made by an individual of known high intelligence? Would such an occurrence be related to an organic impairment, emotional instability, or what?

The Condition-Gestalt might find more practical immediate use in test-batteries constructed specifically for the mentally retarded. It might be administered with the goal in mind of determining the trainability of retarded individuals. One might deduce from present findings that a retarded person who showed "minimal" fluctuation in figures in the various conditions would be a better candidate for certain types of unskilled or semiskilled work than one who showed much fluctuation. Further studies directed along this path would be necessary to determine the correctness of such thinking.

According to Strauss and Kephart (3), brain-injured children can profit from specialized training in the recognition of different perceptual dimensions. The Condition-Gestalt, requiring as it does the reproduction of a wide variety of geometrical figures, might prove useful as a perceptual training device along the lines suggested by Strauss and Kephart. Teachers of the retarded may wish to alter the character of the conditions under which the test series is administered for the purpose of stimulating and maintaining interest in the task.

E. SUMMARY

Groups of retarded and nonretarded subjects were compared as to their performance in a series of conditions on a test-instrument composed of geometrical figures devised by the writer. The physical length of the figures reproduced by the subjects on paper was the single variable explored. The findings may be summarized as follows:

1. The nonretarded group reproduced significantly longer figures, regardless of the condition in which they were drawn, which more closely approximated the actual length of the figures shown.
2. Reproduction of the figures from memory appeared to further decrease the length of certain figures drawn by the retarded group, whereas the nonretarded group was not so affected.
3. The retarded and nonretarded groups were about equally consistent in reproducing the same figures at constant lengths in the successive conditions; the nonretarded group, however, tended to reproduce different figures recalled in the memory condition more alike in length than the retarded group.
4. Positional effects and effects having to do with special gestalt characteristics of the figures themselves appeared more strongly in the reproductions of the retarded group as compared with the nonretarded. It was noted that gross distortions in figures were reproduced only by retarded subjects, usually those having mental ages below six years.

The findings of the study were related to the conceptual model which Freud developed of the psychic apparatus as constructed of several systems possessing functional properties of optical systems. On the basis of the findings,

qualitative differences in the psychic systems of retarded and nonretarded individuals were described.

REFERENCES

1. BENDER, L. A Visual Motor Gestalt Test and Its Clinical Use. New York: American Orthopsychiatric Association, 1938.
2. FREUD, S. The Interpretation of Dreams. New York: Carlton House, 1931.
3. STRAUSS, A., & KEPHART, N. Psychopathology and Education of the Brain-Injured Child, Volume II. New York: Grune & Stratton, 1955.
4. ZUK, G. H. Condition-analysis: A research methodology. *J. Genet. Psychol.*, 1957, **91**, 143-145.

St. Christopher's Hospital for Children
Lawrence and Huntington Streets
Philadelphia 33, Pennsylvania

The Journal of Genetic Psychology, 1959, **94**, 131-141.

INTELLECTUAL APPRAISAL OF PHYSICALLY HANDICAPPED CHILDREN*

Department of Educational Psychology, University of Nebraska

JOSEPH L. FRENCH

Psychologists associated with programs for physically handicapped children continue to be faced with the problem of assessing mental ability. Newland feels that "if we are to plan intelligently for these (physically handicapped) children, such planning has to be in terms of their intellectual potential" (17, p. 11).

Many psychologists have expressed concern with the difference encountered in using available tests in their present form with physically handicapped children. Although the problem is less acute for children over six than at younger ages, it is, nevertheless, present to some extent. In summarizing five recent studies, Strother (21, p. 11) concluded "that in ages over six, standard tests can be applied in the majority of cases." He also added that even with those over six "this does not imply that the results are necessarily reliable or valid measures of intelligence in these cases but merely that the child's physical condition does not preclude the administration of these tests." In some of his own work he considered the validity of the results questionable in nearly one third of the "testable cases." He concluded that great difficulty is encountered with cases who are untestable with the standard tests presently available.

Approximately 15 per cent of 992 cerebral palsied children residing in New Jersey presented extreme testing problems to Hopkins, Bice, and Colton (12, p. 51) in that they could make only the most limited response to any test material. "By the movement of eyes, or perhaps a hand, some of these children were able to indicate their answers to certain multiple choice questions; . . ."

In Buffalo Miller and Rosenfeld found that 20 per cent of 330 children in their study of Cerebral Palsied children could not "be evaluated by the psychologic examination in its standard form" (16, p. 614).

Newland (17, pp. 59-60) discussed the importance of the early location of physically handicapped children in a symposium sponsored by the American Psychological Association, Division of School Psychologists, and the National Society for Crippled Children and Adults:

---

*Received in the Editorial Office on December 18, 1957.

> Granted that some of these handicapped persons are difficult to diagnose early in their lives, we need to keep striving to get them as early as we possibly can in order that appropriate medical and psychological treatment can be supplied at what appears to be a very crucial time. Educational work, in the form of at least socialized experience, can well start as early as two or three years of age. It is encouraging to note that recent legislative provisions for the education of various kinds of handicapped children are tending more and more to make possible the extension of such facilities below the usual lower limit of five or six years.

Later in the presentation (17, p. 64) he said:

> Our literature has been replete with statements as to how various existing devices, or parts of devices, are being utilized in the attempt to assess the intellectual capabilities of these children. I do this, too. But the fact still remains that the results of such efforts are still judgmental and highly subjective. We tend to go thus from the known to the unknown. The less critical (perhaps the more gullible) go so far as to try to pinpoint by *IQ*'s a characterization of behavior that still should be dealt with only in terms of grosser categorizations.

Some of the reasons why tests for physically handicapped children are needed have been set forth by Haeussermann (9, p. 3): They are:

> (*a*) The percentage of retarded children is larger among the cerebral palsied than among the average population;
>
> (*b*) The exploration of possibly existing gaps and deviations in the mental development prior to planning individual training programs results in more complete correction of uneven development;
>
> (*c*) The scarcity of existing training centers calls for careful screening of the available candidates;
>
> (*d*) Long term planning for retarded and defective cerebral palsied children can be undertaken by parents, if reliable tests can be depended upon to give valid results;
>
> (*e*) Premature discouragement of parents and others concerned with the child can be dispelled when it is possible to probe existing abilities in even the most severely involved child by suitable tests.

She pointed out that currently available tests require a verbal or manipulative response by the subject, and she felt that test items for physically handicapped children have to "be constructed in such a way that the child's ability to comprehend an increasing difficulty or abstractness can be observed by the person testing him" (9, p. 4). She feels that the responses to these items should be such that a "simple level of affirmation and negation can serve throughout the whole scale." If this condition can be effected, the burden of proof is shifted from the testee to the item which tests the child's level of comprehension.

Her own (unpublished) scale consists of concrete life-size objects, large clear pictures of objects and children, large pegs of wood matched in pairs of different colors, toy household articles, miniature toys, a sandpaper block and a zwieback, and several other items. In this test the same value is given to a variety of responses in consideration of the severity of the handicap.

The following tests have been selected by the staff of the United Cerebral Palsy Clinic of Miami (1, p. 11) as best meeting the needs of the clinic's clientele and the purposes of the evaluation and treatment team: Ammons Full Range Picture Vocabulary Test, Columbia Mental Maturity Scale, Raven's Progressive Matrices, Leiter International Performance Scale, Vineland Social Maturity Scale, Gesell Preliminary Behavior Inventory, Cattell Infant Intelligence Scale, and the Revised Stanford-Binet, Form *L*. These tests, along with Wechsler's scales, seem to be the most frequently used instruments for determining the intellectual ability of physically handicapped individuals.

To date most clinicians throughout the country have attempted to utilize existing tests for appraising the intelligence of physically handicapped children. Blum, Burgemeister, and Lorge (5, p. 175) concluded that since verbal, motor, and social items are inherent in most currently available standardized tests, "the evaluation of the more severely handicapped cerebral palsied children too often tends to lead to biased estimates." They felt that the Ammons, Matrices, and Columbia represent some efforts to "meet the need for a test of intelligence with items better suited to the limitations of the (young) physically handicapped child so that a more reliable estimate of his mental ability can be obtained" (5, p. 175).

The Ammons Full Range Picture Vocabulary Test (2, 3) consists of a list of 85 words (for two forms) and 16 cards with four line drawings on each. The subject is asked to indicate which one of the four pictures on a plate defines a given word. Several words may be used with the same stimulus card. The norms range in mental age from 2.5 to 16.5 years and adult norms are expressed in percentiles. The subject must comprehend speech or be able to read and have near normal vision to take the test.

Raven's Progressive Matrices (1947), Sets *A*, *Ab*, and *B* (18) are sometimes referred to as the color progressive matrices. These materials can be used with children five to 11. Sets I and II are designed for individuals of above average intelligence and of at least 11 years of age. Norms are provided in terms of expectancies for each age level in terms of percentile rank and "grade" classification (95 per cent and over for Grade I, and five per cent and under for Grade V).

In the administration of the Raven's, a large pattern or matrix from which a portion has been omitted is shown to the subject. The task is to select from six different inserts printed underneath each matrix one which will complete it. A formboard form is recommended for the preschool population, but this form involves more manipulation than the other edition. The directions for the Progressive Matrices can be given in pantomime, but this practice is a deviation from the standardization procedure. Vision must be adequate.

The Columbia Mental Maturity Scale (6) comprises 100 items, each consisting of a set of three, four, or five drawings printed on a six by 19 inch varicolored card. The task of the subject is to select from a series of drawings the one which is unrelated to the others in the series. The mental age norms range from two years, five months to 16 years, three months. The subject should be able to comprehend verbal directions but some pantomime can be used with the more intelligent hearing-impaired subject. As in the case of pantomiming the directions for the Matrices, this procedure is in violation of the standardization technique. Vision must be adequate for the subject to see the various items.

Each of the foregoing tests is a power test and responses may be made by pointing so that verbalization is not necessary. Each test requires near normal vision. Hearing is necessary if the standardization procedures are to be followed. The Ammons and Columbia can be used with a few two and many three-year-old subjects but the Matrices' booklet form cannot be used with subjects under the age of five.

The groups used in standardizing each of these tests are not representative of the entire United States. The Ammons preschool sample is composed of 30 boys and girls at each age level residing in Denver and in the surrounding agricultural areas. Raven worked in England and the Columbia was developed in New York City.

The Columbia Mental Maturity Scale is the most recent of the three tests to become available commercially. The authors state that the test is particularly suitable for use with subjects having serious verbal or motor impairment. However, in the three published studies (13, 8, 7) evaluating this scale when used with physically handicapped and/or mentally retarded children, the general conclusion has been to the effect that the "Stanford-Binet, while still far from a perfect instrument is a better indicator of the level of academic ability" (13, p. 178); that the "results indicate the need for extreme caution in assuming the comparability of test scores" of the Columbia, Stanford-Binet, and the Leiter (8, p. 71); and that it is not "wise to use this scale independently in its present form" (7, p. 132).

The publishers are "in the process of revising and obtaining additional normative data for the Columbia Scale and are acquiring information on its relation to Stanford-Binet and other tests that may modify somewhat the reservations . . . expressed concerning the Columbia Scale" (15).

The literature has been more favorable toward the Ammons and Matrices than toward the Columbia. Taibl recorded the feasibility of administering the Matrices to cerebral palsied subjects six and older in 1951 (22). As Strother pointed out, however, this does not imply that the results are necessarily reliable or valid measures of intelligence in these cases but merely that the child's physical condition did not preclude the administration of these tests.

Richardson and Kobler (19, p. 108) indicated that the Ammons and the Matrices can be used either alone or in combination as may be dictated by the clinical situation. They concluded that the correlations with the Stanford-Binet "appear sufficiently high to employ either of these tests separately, with a preference for the Ammons, the most acceptable procedure in view of the results would be to give each cerebral palsied child both of these appropriate tests."

Allen and Collins advocate using the Ammons, Matrices, and Columbia with severely handicapped children but suggest redrawing the pictures "on very large cards with the items sufficiently separated from each other as to leave little doubt of the indicated response" (1, p. 12). They also suggest a less expensive method—making "use of the child's ability to wink his eyes, nod his head, make some movement with any visible part of the body or to vocalize" (1, p. 13).

Holden (11, p. 95) objected to the Matrices "since it is based on visual perception which is often seriously impaired in brain injured individuals" but felt that the Ammons test "may prove useful in cases of serious physical disability." He has not subjected either of these statements to experimentation.

In working with more complex tests, Sievers and Norman (20, p. 89) used Gessell's Infant Intelligence Scale, the Stanford-Binet, and the Wechsler Bellevue 1 with various age groups in a study involving subjects ranging from nine months to 34 years of age. Some of the causes of low ratings on the Gessell seemed to be "physical handicap, mental defect, poor environment, parental over-protection, and slow development." When the Stanford-Binet was used these additional causes were intimated: "shortness of attention span, perceptual difficulties, and limited speech . . ." Only the verbal part of the Wechsler was used because the adult "subjects could not handle the performance scale."

Richardson and Kohler (19, p. 101) stated that:

> The literature indicates the limitations involved in the use of the more widely accepted tests for such neuro-muscular handicapped children. Despite these limitations, many if not most, research studies as well as clinical centers still employ these tests.

They felt that the researchers using older tests place greater faith in the older tests; or believe that certain modifications, despite standardization procedures, are satisfactory if not desirable.

Several examiners have described modifications of administration procedures for using currently available tests with physically handicapped children. Frequent reference in the literature is made to blinking the eyes in a "yes" or "no" code, shaking the head, and pointing or pushing with hands or feet as means of responding to a multiple choice arrangement of some sort. Still others suggest adjusting the time requirements.

Jewell and Wursten (14, p. 632) found a severely handicapped 12-year-old with athetoid movements of tongue, mouth, and face, who could pass vocabulary items at age level by demonstrating answers with gestures. They feel that a typical technique for coping with manual handicaps can be illustrated by the block design subtests. Blocks are placed before the subject in random fashion on a large piece of paper. Each block is designated by a number on the paper. The subject then verbalizes "start by putting block number 12 on space *A*. By selecting the right blocks, no manipulation of blocks is involved but intelligible speech is necessary." To combat bead-stringing difficulties they recommend (14, p. 633) placing "an open book in front of the subject in such a way that beads slide down the defile from either side to the center. The only problem the child had to really cope with was to get the beads into the proper order by whatever means was available." They tested one subject who picked up the beads with her teeth.

In administering the Porteus Mazes and the Leiter International Performance Scale, Arnold (4) suggested moves to the subject by her own demonstration. If the subject thought the move appropriate, an affirmative response was made.

Miller and Rosenfeld (16, p. 614) used enlarged objects, because they felt the same objects in miniature were not stimulating enough.

In all the suggested adaptations the emphasis is placed on ascertaining whether or not the child has the intellectual grasp of the problem rather than whether the child has the motor skills to carry out the task in a manner standardized for non-handicapped children. No attempt has been made to measure the relative contributions of intellectual grasp and motor skill in determining the position of an item requiring each on a scale. Jewell and Wursten (14, pp. 633-634) observed that these adaptations and deviations from prescribed procedure frequently demanded an ingenuity in getting at the child's comprehension that would be used, perhaps with one child only, and would never be needed again.

Unfortunately, in only one study has an attempt been made to determine how normal children would be scored on these items if they were asked to respond in the same fashion. The one exception is found in a study by Wamba and Marzolf (23, pp. 405-407). Marzolf felt that the subjects did not need to respond either by manipulation or verbalization, or even by indicating an agreement or disagreement. He noted that even with severely handicapped subjects "it is seldom that the use of extra-ocular muscles is so impaired as to make coördinate eye movements unrecognizable. Even when nystagmus exists it is usually possible to discern the locus of regard."

To make use of eye fixations as a means of indicating choice of response to test situations, a box was designed and constructed by Wamba which permits the examiner to observe the subject's eye movements when instructed to look at his choice of correct test-item response.

Raven's Progressive Matrices was chosen as a test to determine whether it is reasonable to suppose that a test administered so that the subject responds by looking at the correct answer will yield results which closely approximate those obtained by the standard mode of test administration.

The "problem" portion of each item of the test was cut from the booklet and pasted on a card. The answer foils were likewise cut from the booklet and pasted on a sheet of plastic in a circular arrangement equidistant from one another. The test was administered to 46 children ranging in age from six to 11 years, without physical handicap, who were randomly assigned to one of two experimental groups, *A* and *B*. Group *A* was first given the test in the usual booklet form and approximately one week later the above described adaptation. For Group *B* the order of giving the two forms was reversed. There was no significant difference between scores resulting from the two forms of the test.

It is recognized that these data are limited in scope; but if Wamba and Marzolf may assume randomness of sample, it appears that tests may be given using eye movements as indicators of response without altering the quantitative findings to a significant degree. These authors concluded that the "use of the procedure is not limited to the Progressive Matrices but may be extended to other kinds of tests, verbal or non-verbal, so long as responses may be indicated by a choice reaction" (23, p. 407).

If eye movements are to be used as indicators of response, the choices must be far enough apart so that it is possible to "discern the locus of regard."

The comments and research results cited in the foregoing paragraphs emphasize the need for the development of an individual test of mental ability that can be used with physically handicapped children.

In considering the desirable characteristics for a new power scale for the appraisal of the mental ability of the physically handicapped, the following points are pertinent: The scale should (*a*) be effective with subjects under six years of age, (*b*) be administered without requiring either speaking or manipulative responses, (*c*) be standardized on a representative sample of a wide area of the United States, (*d*) have several sets of directions and be broad in scope so that in a short time a variety of avenues can be explored.

Heilman has stated that (10, pp. 69-70) "no group of cerebral palsied individuals selected for the standardization group can be considered to be representative of a cerebral palsied universe" and feels that a test standardized on cerebral palsied individuals is meaningless for prediction when used with other cerebral palsied individuals. She points out that even though tests standardized on the blind and the deaf have been useful, the tests for the orthopedically handicapped must be standardized on non-handicapped groups.

To meet the four requirements listed above, the North Central Individual Test of Mental Ability (NCITMA) was developed.

The *NCITMA* includes six point-scale, multiple choice subtests containing from 12 to 30 items similar to those in existing intelligence tests. The subtests are: picture vocabulary, form discrimination, information and comprehension, similarities, size and number, and immediate recall. The test requires neither verbal nor manipulative response by the subject but does require near normal vision and hearing of the child as he is requested verbally to look at the pictorial symbol of the correct answer. Administration of the test requires about 25 minutes.

The sample was secured in Nebraska and Illinois homes and schools closely resembling the community structure and occupational classification characteristics of the North Central Section of the United States as defined by the 1950 census. The *NCITMA* was administered to over 500 physically normal subjects residing in this area to establish the initial evaluation of the test and the preliminary normative data.

To achieve a satisfactory trial administration of the original *NCITMA* for the purpose of refining the items and administrative procedure through a subsequent analysis, a sample of 100 children was selected. To be more certain of the intellectual level of the sample, the Stanford-Binet, Form *L*, was administered to half of these subjects. Entries in the age distribution included the Stanford-Binet mental age in place of the chronological age for these 50 subjects. Since the mental age was used in the pretest sample, further stratification was not attempted. The item analysis of the pretest data was conducted as suggested by Flanagan and the correlation of each item with total test score was determined.

To achieve a satisfactory standardization of the test in its final form, a group of 400 subjects or 100 three-, four-, five-, and six-year-olds was selected. This number was considered sufficient for this purpose inasmuch as considerable attention was given to the representation.

Reliability coefficients of the total test and of the subtests were obtained by the split half technique (Spearman-Brown) and by test and retest for each age level and for all ages combined. When the four ages were combined, a Spearman-Brown coefficient of .965 and a test-retest coefficient of .967 were obtained for the total test. The reliability of each subtest for the four age levels combined was also computed by the split half technique. The correlations of each subtest with each other subtest also demonstrated the reliability of the instrument.

The validities of the subtests and the full test were assessed by several methods. Among those methods the most important evidence of validity was considered to be the .886 correlation coefficient representing the relationship between chronological age and the total score of the test. In addition to this correlation coefficient the demonstration of increasingly high mean scores on each subtest through the chronological ages of the subjects; the differential performance of occupational groups as indicated by the means of each group; and the several correlation coefficients indicating the relationship between the rankings by a teacher of several individuals and the rankings of these individuals by their scores on the *NCITMA* also indicated validity. At each age level each item was correlated with each of the subtests to determine the appropriateness of classifying the items into subtests.

The normative data are expressed in terms of mental age, deviation *IQ*, and percentile rank.

When the *NCITMA* was used with some physically handicapped children, average and even superior *IQ*'s were obtained. These illustrative cases show that the items utilized in the *NCITMA* are not outside the realm of the severely physically handicapped and that the items in the *NCITMA* afford a means of comparing the mental ability of the physically handicapped with the mental ability of the physically normal children in their reaction to these items.

Summary: The current status of appraising the intelligence of physically handicapped children has been presented. The tests currently in use have been identified as have many of the suggested modifications of these tests. This review of research indicated a need for the development of a new power scale for appraising the mental ability of the physically handicapped and consequently the North Central Individual Test of Mental Ability was developed. The initial standardization of the *NCITMA* was abstracted.

REFERENCES

1. ALLEN, R. M., & COLLINS, M. G. Suggestions for the adoptive administration of intelligence tests for those with cerebral palsy. *C. P. Rev.*, 1955, **16**, 11-14.
2. AMMONS, R. B. Full Range Picture Vocabulary Test. Louisville: Southern Universities Press, 1948.
3. AMMONS, R. B., & HOLMES, J. C. The Full Range Picture Vocabulary Test: III. Results for a preschool age population. *Child Devel.*, 1949, **20**, 5-14.
4. ARNOLD, G. F. A technique for measuring the mental ability of the cerebral palsied. *Psychol. Serv. Center J.*, 1951, **3**, 171-178.
5. BLUM, L. H., BURGEMEISTER, B., & LORGE, I. Trends in estimating the mental maturity of the cerebral palsied child. *Except. Child.*, 1951, **17**, 174-177.
6. ———. Columbia Mental Maturity Scale Manual. Yonkers-on-Hudson: World Book, 1954.
7. FRENCH, J. L., & WORCESTER, D. A. A critical study of the Columbia Mental Maturity Scale. *Except. Child.*, 1956, **23**, 111-113; 131-132.
8. GALLAGHER, J. J., BENOIT, E. P., & BOYD, H. F. Measures of intelligence in brain damaged children. *J. Clin. Psychol.*, 1956, **12**, 69-72.
9. HAEUSSERMANN, E. Evaluating the developmental level of cerebral palsy preschool children. *J. Genet. Psychol.*, 1952, **80**, 3-23.
10. HEILMAN, A. E. (Discussant of) Educational and vocational planning for the cerebral palsied child (by T. E. Newland). In *Psychological Problems of Cerebral Palsy*. Chicago: National Society for Crippled Children and Adults, 1952. Pages 69-73.
11. HOLDEN, R. H. A review of psychological studies in cerebral palsy: 1947-1952. *Amer. J. Ment. Def.*, 1952, **57**, 92-99.
12. HOPKINS, T. W., BICE, H. V., & COLTON, K. C. Evaluation and Education of the Cerebral Palsied Child. Washington, D. C.: International Council Exceptional Children, 1954.
13. JOHNSON, G. O., NEELY, J. H., & ALLING, R. L. A comparison of the 1937 revision of the Stanford-Binet (Form *L*) and the Columbia Mental Maturity Scale. *Except. Child.*, 1956, **22**, 155-157.
14. JEWELL, B. T., & WURSTEN, H. Observations on the psychological testing of cerebral palsied children. *Amer. J. Ment. Def.*, 1952, **56**, 630-637.
15. LENNON, R. T. Personal communication. October, 1957.
16. MILLER, E., & ROSENFELD, G. B. The psychologic evaluation of children with cerebral palsy and its implications in treatment. *J. Pediat.*, 1952, **41**, 613-621.
17. NEWLAND, T. E. Educational and vocational planning for the cerebral palsied child. *Psychological Problems of Cerebral Palsy*. Chicago: National Society for Crippled Children and Adults, 1952. Pages 54-69.

18. Raven, J. C. Guide to Using Progressive Matrices (1947). London: Lewis, 1951.

19. Richardson, E. J., & Kobler, F. J. Testing the cerebral palsied. *Except. Child.*, 1954, **21**, 101-103; 108-109.

20. Sievers, D. J., & Norman, R. D. Some suggestive results in psychometric testing of the cerebral palsied with Gesell, Binet, and Wechsler scales. *J. Genet. Psychol.*, 1953, **82**, 69-90.

21. Strother, R. The psychological appraisal of children with cerebral palsy. In *Psychological Problems of Cerebral Palsy*. Chicago: National Society for Crippled Children and Adults, 1952, 5-9.

22. Taibl, R. M. An investigation of Raven's "Progressive Matrices" as a tool for the psychological evaluation of cerebral palsied children. (Unpublished doctor's dissertation, Univ. Nebraska, 1951.)

23. Wamba, D. E., & Marzolf, S. S. Use of eye movements as response indicators in testing the physically handicapped. *J. Clin. Psychol.*, 1955, **11**, 405-411.

214 Manor Drive
Columbia, Missouri

The Journal of Genetic Psychology, 1959, **94**, 143-146.

SIX PREDICTORS OF COLLEGE ACHIEVEMENT*

Hofstra College

Harold L. Henderson and Sherman H. Masten[1]

A. Problem

This research considered the relative and combined validities of achievement data from the high school transcript together with college aptitude and scholastic ability tests for predicting the first semester averages for the 1955 freshman class at Hofstra College. The study was also planned as a cross-validation of earlier research (2). Unique features of this investigation were the attention given to reduction in the size of correlation coefficients through restriction in range and the comparison of two multiple correlation methods.

B. Predictors

At the time of application to the College New York State Regents Examination grades are available for most of the students, and the high school academic record is complete except for the last term's grades. These instruments are then administered to all prospective students: the Coöperative Reading Test (C2, Form Y); the Coöperative Effectiveness of Expression Test (Form B2)[2]; The *ACE* Psychological Examination for College Freshmen (Form 1952).

C. Sample

Students used in this study were those 332 (out of an initial accepted and attending freshmen class of 564) for whom *all* these variables were available at the time of completion of the first semester: at least one English Regents grade in per cent; the average of at least three other Regents' scores in both social and natural science subjects; high school grade average in per cent; Coöperative Reading scaled score; Coöperative Effectiveness of Expression scaled score, and *ACE* Psychological Examination total raw score.

---

*Received in the Editorial Office on December 23, 1957.

[1] The authors are indebted to the Deans of Admission and Students, E. T. Thomas and J. M. Watts, for their interest and cooperation in this research, and to Phyllis Richel, graduate student in Psychology, for her statistical assistance.

[2] This test is used primarily for placement in Freshman English after admission.

D. CRITERION

At the end of the first semester of the freshmen year, the average grade in per cent achieved by full-time students completing at least 12 semester hours was used as the criterion against which the predictors were correlated.

E. PROCEDURES

Following the calculations of zero order correlations and intercorrelations, the correlations of those predictors that had been utilized by the Admissions Office in screening, and were consequently reduced in size, were corrected for restriction in range (4).

The Wherry-Doolittle Test Selection procedure (1) was then used with both corrected and uncorrected correlation coefficients. To test the utility of the Jenkins (3) short cut, this method was also employed.

F. RESULTS

Table 1 presents the distribution statistics and correlation coefficients, and shows how the study sample had already been preselected on some of the variables under investigation.

TABLE 1
MEANS, STANDARD DEVIATIONS, VALIDITIES, AND INTERCORRELATIONS OF THE PREDICTORS AND THE CRITERION[a]
(N = 332)

	1	2	3	4	5	6	Applicants for admission[b] Mean	Applicants for admission[b] *SD*	Accepted students Mean	Accepted students *SD*
Criterion:										
First semester average										
"Corrected"[c]	—	59	—	53	45	34	—	—	—	—
"Not corrected"	58	57	45	40	28	22	—	—	76.6	6.8
Predictors:										
1. Regents Avg.[d]		70	40	27	25	28	—	—	77.1	6.6
2. High School Avg.			51	23	20	19	76.9	6.0	79.1	5.7
3. English Regts. Avg.				49	43	36	—	—	75.9	7.4
4. Coop. Read.					58	62	55.2	8.2	58.2	5.7
5. Coop. Effect.						57	54.5	8.4	58.0	5.5
6. *ACE* Total							106.5	24.8	115.5	15.2

[a] For ease in reading correlations the decimal point has been dropped and the correlation rounded to two places.
[b] N equals 1069.
[c] Refers to correction of *r*'s for restriction in range (4) for *only* the correlations of predictors with the criteria. Intercorrelations of predictors are *not* corrected; thus, multiple correlations in later tables are slightly inflated.
[d] Excluding English Regents.

Table 2 lists the ordering of predictors for first semester average, multiple *R*'s and Beta weights, and shows how the correction for restriction in range effects the ordering of predictors.

TABLE 2

MULTIPLE CORRELATIONS AND BETA WEIGHTS

(N = 332)

Predictor variables	R	Betas Six predictors	 Four predictors
(*r*'s not corrected for restriction in range)			
Regents	.580	.314	.322
High School[a]	.622	.267	.295
Coop. Read.	.663	.288	.324
ACE	.673	—.138	—.127
English Regents[b]	.672	.089	—
Coop. Effect.[b]	.668	.019	—
(*r*'s corrected for restriction in range)			
High School	.59	.340	.325
Coop. Read.[a]	.712	.365	.300
Regents	.731	.248	.230
Coop. Effect.	.738	.192	.146
ACE[b]	.739	—.113	—
Eng. Rgts.[b]	.737	—.044	—

[a] The correlation given is the combined *R* for the predictor given on the same line and all predictors previously listed above it.

[b] Naturally these predictors need not be utilized in the regression equation, as they contribute nothing additional.

Table 3 compares the *R*'s obtained from the Wherry-Doolittle and Jenkins' methods, showing from observation (without the need for significance tests) that they produce rather similar results in most instances but substantial differences in a few cases.

TABLE 3

COMPARISON OF WHERRY-DOOLITTLE AND JENKINS MULTIPLE CORRELATIONS[a]

(N = 332)

Predictor variables	*R*'s Doolittle	 Jenkins
Correlations not corrected for restriction in range		
Regents and High School	.622	.625
Regents and High School and Coop. Read.[b]	.663	.636
Rgts., H. S., Coop. R., and *ACE*	.673	.649
Rgts., H. S., Coop. R., *ACE*, and Eng. Rgts.	.672	.695
Rgts., H. S., Coop. R., *ACE*, Eng. R., and Coop. Effect.	.668	.795
H. S. and Coop. R.	.633	.633
H. S., Coop. R., and *ACE*[b]	.634	639
H. S., Coop. R., *ACE*, and Coop. Effect.	.630	.677
Rgts., Coop. R., and Eng. R.	.640	.641
Rgts., Coop. R., Eng. R., and *ACE*[b]	.646	.672
Correlations corrected for restriction in range		
H. S., Coop. R., and Rgts.	.731	.725
H. S., Coop. R., Rgts., Coop. Effect., *ACE*, Eng. R.	.737	.794

[a] Some combinations are not comparable in terms of variables present (which ones and order); these are therefore not listed. The two methods do not always allow the tests to go into the Multiple *R* in the same order.

[b] The correlation given is the combined *R* for the predictor given on the same line and all previously listed predictors.

Table 4 enumerates the multiple *R*'s and Beta weights for "practical prediction" wherein the fewest variables that are always available for all students, and give useful predictions, are employed. Regents scores are not always available for all applicants, whereas high-school average is; thus it is used in preference to the Regents.

TABLE 4
MULTIPLE CORRELATIONS AND BETA WEIGHTS FOR PREDICTORS ALWAYS AVAILABLE
(N = 332)

Predictor variables	R	Betas Four predictors	Betas Two predictors
	(r's not corrected for restriction in range)		
High School	.574	.510	.612
Coop. Read.	.633	.341	.389
ACE[a]	.634	—.131	—
Coop. Effect.[a]	.630	.053	—
	(r's corrected for restriction in range)		Three predictors
High School	.590	.485	.492
Coop. Read.	.712	.358	.331
Coop. Effect.	.721	.193	.176
ACE[a]	.720	—.084	—

[a] These predictors contribute nothing additional.

G. SUMMARY

This research reports the correlations, intercorrelations, and multiple correlations for New York Regents Average excluding English Regents, New York English Regents Average, High School Average, Coöperative Reading score, Coöperative Effectiveness of Expression score, *ACE* Psychological Examination score, and first semester college average for the 1955 Freshman Class at Hofstra College. Effects upon *R*'s of correcting variables for restriction in range, comparison of the Wherry-Doolittle and Jenkins methods for *R*, and selection of the fewest, most useful, predictors are shown.

REFERENCES

1. GARRETT, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1953.
2. HENDERSON, H. L. Prediction of freshmen grades in a Long Island college. *Educ. & Psychol. Meas.* 1957, **17**, 623-627.
3. JENKINS, W. H. An improved short-cut method for multiple *R*. *Educ. & Psychol. Meas.*, 1952, **12**, 316-322.
4. THORNDIKE, R. L. Personnel Selection. New York: Wiley, 1949.

Hofstra College
Hempstead, New York

The Journal of Genetic Psychology, 1959, **94**, 147-149.

BOOKS

Now that there is a special *APA* journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a secondary order of publication unless they carry information that is equally as important as the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

BOOKS RECENTLY PUBLISHED

Ainsworth, M. D., & Ainsworth, L. H. Measuring Security in Personal Adjustment. Toronto: Univ. Toronto Press, 1958. Pp. 98.

Atkinson, J. W. (*Ed.*). Motives in Fantasy, Action, and Society. Princeton: Van Nostrand, 1958. Pp. 873.

Ausubel, D. P. Theory and Problems of Child Development. New York: Grune & Stratton, 1958. Pp. 650.

Banks, F. Teach Them to Live. New York: International Univ. Press, 1958. Pp. 287.

Bartley, S. H. Principles of Perception. New York: Harper, 1958. Pp. 482.

Bay, C. The Structure of Freedom. Stanford: Stanford Univ. Press, 1958. Pp. 419.

Bernert, E. H. America's Children. New York: Wiley, 1958. Pp. 185.

Bertocci, P. A. Religion as Creative Insecurity. New York: Association Press, 1958. Pp. 128.

Beutel, F. K. Some Potentialities of Experimental Jurisprudence as a New Branch of Social Science. Lincoln: Univ. Nebraska Press, 1957. Pp. 440.

Bossard, J. H. S., & Boll, E. S. Why Marriages Go Wrong. New York: Ronald Press, 1958. Pp. 224.

Broadbent, D. E. Perception and Communication. New York: Pergamon, 1958. Pp. 338.

Campbell, J. (*Ed.*). Man and Time. New York: Pantheon, 1957. Pp. 414.

Campbell, R. F., Corbally, J. E., Jr., & Ramseyer, J. A. Introduction to Educational Administration. Boston: Allyn & Bacon, 1958. Pp. 434.

Cantril, H. The Politics of Despair. New York: Basic Books, 1958. Pp. 269.

Chisholm, B. Prescription for Survival. New York: Columbia Univ. Press, 1957. Pp. 92.

Christensen, H. T. Marriage Analysis. New York: Ronald Press, 1958. Pp. 645.

Coelho, G. V. Changing Images of America. Glencoe: Free Press, 1958. Pp. 145.

Cohen, J. Humanistic Psychology. London: Allen & Unwin, 1958. Pp. 206.

Cohen, N. E. Social Work. New York: Henry Holt (Dryden), 1958. Pp. 404.

Culin, S. Games of the Orient. Rutland: Tuttle, 1958. Pp. 177.

Daugherty, W. E., & Janowitz, M. A Psychological Warfare Casebook. Baltimore: Johns Hopkins Press, 1958. Pp. 880.

Eissler, R. S., Freud, A., Hartmann, H., & Kris, E. (*Eds.*). The Psychoanalytic Study of the Child. Vol. XII. New York: International Univ. Press, 1958. Pp. 417.

Farnsworth, P. R., & McNemar, Q. Annual Review of Psychology. Vol. 9. Palo Alto: Annual Reviews, 1958. Pp. 543.

Firth, R. We, The Tikopia. London: Allen & Unwin, 1958. Pp. 605.

Galdston, I. (*Ed.*). Panic and Morale. New York: International Univ. Press, 1958. Pp. 340.

Goody, J. (*Ed.*). The Developmental Cycle in Domestic Groups. New York: Cambridge Univ. Press, 1958. Pp. 145.

Gordon, M. M. Social Class in American Sociology. Durham: Duke Univ. Press, 1958. Pp. 281.

Greene, T. M. Moral, Aesthetic, and Religious Insight. New Brunswick: Rutgers Univ. Press, 1958. Pp. 141.

Gruenberg, S. M. The Parents' Guide to Everyday Problems of Boys and Girls. New York: Random House, 1958. Pp. 363.

Heider, F. The Psychology of Interpersonal Relations. New York: Wiley, 1958. Pp. 322.

Jung, C. G. The Undiscovered Self. Boston: Little, Brown, 1958. Pp. 113.

Kemp, C. F. The Church: The Gifted and the Retarded Child. St. Louis: Bethany Press, 1958. Pp. 189.

Lesser, S. O. Fiction and the Unconscious. Boston: Beacon, 1957. Pp. 322.

Lindzey, G. (*Ed.*). Assessment of Human Motives. New York: Rinehart, 1958. Pp. 273.

Luszki, M. B. Interdisciplinary Team Research: Methods and Problems. New York: New York Univ. Press, 1958. Pp. 355.

Mannheim, K. Systematic Sociology. New York: Philosophical, 1958. Pp. 169.

McClelland, D. C., Baldwin, A. L., Bronfenbrenner, U., & Strodtbeck, F. L. Talent and Society. Princeton: Van Nostrand, 1958. Pp. 275.

McKercher, P. The Bomb in Your Life. New York: Vantage, 1958. Pp. 89.

Muensterberger, W., & Axelrad, S. (*Eds.*) Psychoanalysis and the Social Sciences. New York: International Univ. Press, 1958. Pp. 297.

Newby, H. A. Audiology: Principles and Practice. New York: Appleton-Century-Crofts, 1958. Pp. 342.

Nye, F. I. Family Relationships and Delinquent Behavior. New York: Wiley, 1958. Pp. 168.

Podolsky, E. (*Ed.*). The Neuroses and Their Treatment. New York: Philosophical, 1958. Pp. 555.

Saul, L. J. Technic and Practice of Psychoanalysis. Philadelphia: Lippincott, 1958. Pp. 244.

Sawrey, J. M., & Telford, C. W. Educational Psychology. Boston: Allyn & Bacon, 1958. Pp. 498.

Schutz, W. C. Firo. New York: Rinehart, 1958. Pp. 267.

Schwarz, B. E., & Ruggieri, B. A. Parent-Child Tensions. Philadelphia: Lippincott, 1958. Pp. 238.

Scott, J. P. Aggression. Chicago: Univ. Chicago Press, 1958. Pp. 149.

Sechrest, C. A. New Dimensions in Counseling Students. New York: Columbia Univ., 1958. Pp. 119.

Service, E. R. A Profile of Primitive Culture. New York: Harper, 1958. Pp. 474.

Sheldon, H. D. The Older Population of the United States. New York: Wiley, 1958. Pp. 223.

Stein, H. D., & Cloward, R. A. (*Eds.*). Social Perspectives on Behavior. Glencoe: Free Press, 1958. Pp. 666.

Tagiuri, R., & Petrullo, L. (*Eds.*). Person Perception and Interpersonal Behavior. Stanford: Stanford Univ. Press, 1958. Pp. 390.

Tinker, M. A., & Russell, W. A. Introduction to Methods in Experimental Psychology. New York: Appleton-Century-Crofts, 1958. Pp. 282.

Tolman, E. C. Behavior and Psychological Man. Berkeley: Univ. California Press, 1958. Pp. 269.

Vander Veldt, J. H., & Odenwald, R. P. Psychiatry and Catholicism. New York: McGraw-Hill, 1957. Pp. 474.

Walker, C. R. Toward the Automatic Factory. New Haven: Yale Univ. Press, 1957. Pp. 232.

Watson, J. B. Behaviorism. Chicago: Univ. Chicago Press, 1958. Pp. 308.

Welford, A. T. Ageing and Human Skill. New York: Oxford Univ. Press, 1958. Pp. 300.

WHEELIS, A. The Quest for Identity. New York: Norton, 1958. Pp. 250.

WHITE, V. Studying the Individual Pupil. New York: Harper, 1958. Pp. 238.

WIKLER, A. The Relation of Psychiatry to Pharmacology. Baltimore: Williams & Wilkins, 1957. Pp. 322.

WILLEY, R. D., & STRONG, W. M. Group Procedures in Guidance. New York: Harper, 1957. Pp. 548.

WINCH, R. F. Mate-Selection: A Study of Complementary Needs. New York: Harper, 1958. Pp. 349.

WINNICOTT, D. W. The Child and the Outside World: Studies in Developing Relationships. New York: Basic Books, 1957. Pp. 190.

WOLFFLIN, H. The Sense of Form in Art. New York: Chelsea, 1958. Pp. 230.

WOLPE, J. Psychotherapy by Reciprocal Inhibition. Stanford: Stanford Univ. Press, 1958. Pp. 239.

WOODWORTH, R. S. Dynamics of Behavior. New York: Henry Holt, 1958. Pp. 403.

WORMHOUDT, A. Hamlet's Mouse Trap. New York: Philosophical, 1956. Pp. 221.

YATES, D. H. Psychology You Can Use. New York: Crowell, 1957. Pp. 248.

YOUNG, R. (*Ed.*). Approaches to the Study of Politics. Evanston: Northwestern Univ. Press, 1958. Pp. 382.

ZAEHNER, R. C. Mysticism Sacred and Profane. New York: Oxford Univ. Press, 1957. Pp. 256.

$14.00 per year
$10.00 per volume
Single Numbers $5.00

QUARTERLY
Two volumes per year

June, 1959
Volume 94, Second Half

Founded by G. Stanley Hall in 1891, reorganized by Carl Murchison in 1925

THE JOURNAL OF GENETIC PSYCHOLOGY

Child Behavior, Animal Behavior, and Comparative Psychology

If this space should be unstamped, this is the regular library edition. But if this space is stamped with a designating title, this is a special edition, sold

under the restrictions of a liberal contract, and may not be resold for a period of five years from the date of publication.

JUNE, 1959

(Manuscripts are printed in the order of final acceptance)

Self-confidence, ability, and the interest-value of tasks 153
By Joseph F. Rychlak

Comparing the effectiveness of the ego-strength Q-sort test by use of R- and Q-methodologies 161
By Russell N. Cassel

A developmental analysis of proverb interpretations 169
By Claudia Richardson and Joseph Church

Anxiety reduction and novelty as goals of visual exploration by monkeys . 181
By David Symmes

Behavior of the opossum in the Fink Arrow Maze 199
By W. T. James

(OVER)

Copyright, 1959, by The Journal Press

Provincetown, Massachusetts, U. S. A.

Entered [illegible]ond-class matter May 11, 1937, at the post-office at
Pro[illegible]wn, Mass., under the Act of March 3, 1879

The effects of long-term hospitalization or institutionalization upon the language development of children 205
By Arthur D. Haggerty

The relation between pre-irradiation volitional activity and post-irradiation death order in the rat 211
By Arnold A. McDowell and W. Lynn Brown

Parallel observations on avoidance behavior and endocrine response under reserpine 215
By Stephen B. Coslett and Alfred B. Shaklee

Conceptual block sorting as a function of number, pattern of assignment, and strength of labeling responses 221
By Harvey Lacey and Albert E. Goss

Infantile illness and subsequent mental and emotional development . . . 233
By D. H. Stott

Immanent justice in children: A review of the literature and additional data 253
By Gene R. Medinnus

The influence of auditory stimulation on aniseikonic perception: II and III. Business school women and college women 263
By James E. Simpson

Rorschach indices of developmental level 267
By Leslie Phillips, Stanley Kaden, and Marvin Waldman

The function of verbal labels in the discrimination of subtle stimulus differences 287
By Robert S. Albert

BOOKS 297

BOOKS RECENTLY PUBLISHED 297

The Journal of Genetic Psychology, 1959, **94**, 153-159.

SELF-CONFIDENCE, ABILITY, AND THE INTEREST-VALUE OF TASKS*

Department of Psychology, Washington State College

JOSEPH F. RYCHLAK

A. PROBLEM

Many studies (cited in references) have attempted to identify the relationship between interests and abilities. Although the results have been inconclusive, there are indications that interest value or task attractiveness may be related to estimated rather than actual ability. Cartwright (2) found that subjects (*S*s) who were told they had failed a task (jigsaw puzzle, etc.) were more likely to see that task as less attractive than previously rated. However, Cartwright made a further tentative conclusion on the basis of his data, which was later supported by Gebhard (6). This general hypothesis stated that the attractiveness of an activity is determined not only by past experience of success or failure but by the expectation of future success or failure as well. "This fact suggests that a person's view of his future is the primarily decisive factor in producing changes in attractiveness . . . Past failure does not always lead to an expectation of future failure" (2, p. 13).

This suggests a personality construct—which might be termed self-confidence—coming into play to influence an *S*'s typical self-evaluation. Thus, at a certain time in his development a child's confidence level is fairly well established regarding various activities. Should he view himself as lacking in numerical ability, he will rate the interest value of a forthcoming arithmetical task according to the expected failure, regardless of how well he might actually do in a subsequent attempt. On the other hand, the more self-confidence a child has in any given area, the less likely is he to be influenced in his ratings of interest by more recent failures in that area.

An important consideration here is the degree of feedback an *S* has concerning his ability, or performance level. Using a design which did not include giving the *S* a score (no feedback), Hirsch (7) noted the following: Rankings of self-estimated ability and interest (preference) correlated to a higher degree (+ .48) than either self-estimated ability and actual ability (+ .39) or actual ability and interest (+ .29). It is important to note that

---

* Received in the Editorial Office on January 9, 1958.

Hirsch's *S*s were juvenile delinquents, and therefore presumably not school achievers. Thus, when *S*s with low confidence in school-like tasks are not told how well they have actually done on such tasks, they tend to estimate their relative performance in terms of the interest value the tasks held for them. Whether this would be true of school achievers (i.e., *S*s with high confidence) has not yet been established.

The primary aim of the present study was to test the following hypothesis: If no feedback (or very little feedback) on actual performance is given, separate rankings of task interest value and self-estimated ability will show a greater relationship in low than high confidence groups. If all *S*s ranked their task interests according to their expected successes on the tasks, then the more confident *S* should see several, if not all of the tasks as potential success experiences. In facing a unique series of tasks this suggested a tentative interest ranking, since any test could prove the one easiest to master. Because he is used to success, and not told about his score, the confident *S* would also perceive more successes when taking the tasks. The *S* with little confidence would expect fewer successes and perceive fewer successes following a performance in which he was not *told* differently. Consequently, after completion of all tasks, an ability ranking made by the confident *S* should bear less relation to initial interest than the comparable rankings made by an *S* with little confidence, who ranks along a more stable continuum of expected ability. Other aims of this research included the relationship of confidence level to actual ability levels, the accuracy of estimated ability, and the change in task interest as a function of experience in the tasks.

B. METHOD

1. *Tasks*

The tasks were four of the performance subtests from the Wechsler Intelligence Scale for Children (*WISC*) (11). They were introduced and administered to the *S*s in the following order: Block Design, Picture Completion, Mazes, and Picture Arrangement. No other subtests were administered, and the same experimenter (*E*) did all of the testing.

2. *Subjects*

Subjects were 60 boys chosen from an elementary school located in a lower middle-class neighborhood.[1] Grades 3 through 5 were used, with the

---

[1] The author would like to express his appreciation to Dr. Leonard O. Andrews of The Ohio State University and Miss Imogene Cole of the Columbus Ohio Public School System for their invaluable coöperation in helping to arrange subjects and testing rooms for the present research.

*S*s ranging in age from 8 to 11 years. As a control measure, only those boys were used as *S*s who had scored within the Average range (i.e., 90 to 110) on a paper-pencil type *IQ* test (Kuhlmann-Anderson). All of the boys within the Average range in each class were included, and none of the *S*s had taken a *WISC* previous to this study. Because of the greater number of children in the lower grades, about two-thirds of the entire sample was drawn from Grades 3 and 4. Although their inclusion did not alter the findings, six boys tested from the 6th grade were dropped from the study because they were not sufficient in number to be representative of this age.

3. *Procedure*

Subjects made three rankings of the four subtests: two of preference or interest value and one of relative ability. At the outset of the experimental session, after being introduced to the tasks, *S*s ranked them as to how "interesting" they seemed, and at the very end of the session re-ranked their preferences on the basis of which subtests they would "most like to do over again, just for the fun of it." Both of these preference rankings are viewed as reflecting interest value. Coming between the two rankings of preference was one ranking of self-estimated ability, done on the basis of estimated high to low score. The experimental hypothesis was that *S*s with low confidence would show a greater relationship between the *initial* preference ranking and the ranking of self-estimated ability than would *S*s with high confidence.

The specific procedure was as follows: After entering the experimental room each *S* was directed to a low table, on which lay the four subtests in the order of their administration. The *S* was informed that he would be taking the tasks and then each was demonstrated according to Wechsler's standard instructions. The *S* did not perform the tasks at this time, but merely observed the *E*'s demonstration of what eventually would be required. At the completion of all four introductions the *S* was routinely asked if these tasks seemed interesting to him. In every instance the boy responded in the affirmative. He was then asked to point to the most interesting task, following with the second, third, and fourth tasks in the order of their interest. This constituted the ranking of initial interest.

After this initial ranking the *S* was asked to take a seat at a nearby desk, and the four subtests were administered. Following completion of each subtest the *S* was asked to rate his performance on the task just finished as either not good, fair, or very good. These ratings were taken as measures of confidence. The *E* took special precaution not to reward a *S*'s performance through words of praise, etc. Except for the standard instructions, the

E remained silent and expressionless. Thus, the confidence ratings depended entirely upon the *S*'s personal feelings of success or failure in the task at hand. He had no extra-personal standard by which to evaluate his performance.

After he had rated his performance on the last task, the *S* was asked to return to the low table where the subtests were arranged in the same order as at the outset of the session. He was asked to identify the task on which he felt he had obtained his highest score, following with the second, third, and fourth scores. This constituted his self-estimated ability ranking. Finally, as the second ranking of interest, each *S* was asked to point to the subtest he would most like to do over again, just for the fun of it, following with the second, third, and fourth as before.

There was good reason to expect a halo-effect between the ranking of estimated ability and the second ranking of task interest, coming as they did one after the other. This was of no importance to the major experimental hypothesis, however, since it referred to the *initial* interest ranking and the estimated ability ranking. Indeed, it seemed significant to determine whether differences in halo-effect might result between high and low confidence groups.

To distinguish highs and lows the confidence ratings were numbered as follows: not good received 1-point, fair received 2-points, and very good was given 3-points. This made the range of total confidence scores 4 to 12—the more points the greater the assumed confidence. On the basis of these total scores each school grade was divided into equal halves, giving a total for the three grades of two confidence groups (titled low and high, although these are not extreme groups) with 30 *S*s each. The median confidence rating for the lows was 6.35 (*MD* .43) and for the highs 9.00 (*MD* .56).

The reader is cautioned not to confuse the confidence *ratings* (independent variable) with the estimated ability *rankings* (dependent variable). It is true that both are a result of a self-evaluation regarding ability. However, the *E* could not legitimately rank the tests *for* the *S* on the basis of the confidence ratings. Each *S* had at least two tests with the same rating, and most often two ties in a ranking of four would have resulted. More importantly, when a *S* ranks his performance on four tests he has the benefit of weighing one performance against all of the rest. This was not true of the ratings. Indeed, the first confidence rating had no possible comparison since the *S* had not yet taken the second test. The second test rating had only the first as a rough comparison, etc. Whereas the confidence ratings reflect the *S*'s evaluation of his performance in tasks of this nature, the estimated ability rankings reflect the *S*'s perceived strong and weak points relative to the fact that he felt his general performance was "very good" or "not good."

C. RESULTS

The subtests were scored and pro-rated according to Wechsler's instructions (11). Resulting mean Performance *IQ*s were 101.84 for the high confidence group and 97.82 for the low confidence group. This difference failed to reach the .10 level of significance ($t = 1.481$) (8), consistent with the belief that although the *S*s of this experiment were roughly equal in ability, they tended to evaluate their abilities differently because of personality factors.

In order to demonstrate the extent of relationship between the three rankings of the study, rhos (8) were calculated. Preliminary to the tests of significance, Difference scores (*D*-scores) were determined for each *S* by totaling the differences in any two of his rank orderings of the four subtests. Following scoring, the *E* arranged the subtests according to their Wechsler weighted scores so that *D*-scores could be determined between the estimated ability ranking and the actual ability ranking (i.e., how well the *S* actually scored on each subtest, relative to the other three). Since the *D*-scores did not distribute themselves normally, chi-square was chosen as the method of testing significance. For each comparison the median *D*-score for the *combined* confidence groups was determined, and then the *S*s in each confidence group dichotomized as falling above or below this value. A 2 × 2 table was arrayed according to the following: high and low confidence; above and below the median *D*-score of the combined groups. Table 1 displays the median rhos, median *D*-scores, and chi-square tests of significance for the

TABLE 1
MEDIAN RHOS, MEDIAN *D*-SCORES AND CHI-SQUARE TESTS OF SIGNIFICANCE

Rankings compared	Median rho of confidence gp.		Median *D*-score of confidence gp.*		Chi-square of *D*-scores
	High	Low	High	Low	
Initial Interest & Self-est. Ability	—.30	.45	5.63	3.62	25.11**
Self-est. Ability & Actual Ability	.12	.26	4.83	4.49	0.28
Initial & Final Interest	.30	.48	3.82	3.30	1.84
Self-est. Ability & Final Interest	.50	.78	3.30	1.62	2.40***

* These are the median *D*-scores for the individual groups, not to be confused with the median *D*-score of the combined groups used in the chi-squares.
** Significance at the .001 level requires chi-square of 10.83 (8).
*** Significance at the .10 level requires chi-square of 2.71 (8).

high and low confidence groups, according to the various rankings under comparison.

As a rough check on ranking reliability between groups (though not completely legitimate, ref. above), the *E* ranked all tasks on the basis of the confidence ratings (1, 2, or 3 points). Several ties resulted, but *D*-scores were calculated between these pseudo-estimated ability rankings and the *S*'s own ability ranking. A 2 × 2 table was arranged similar to the others of the study. This chi-square failed to reach significance (1.38), suggesting that the experimental differences cannot be accounted for on the basis of "ranking" unreliability.

D. DISCUSSION

Finding that the groups failed to differ significantly in pro-rated *IQ* holds much interest for the psychologist, since it supports his claim that personality factors may come into play resulting in self-evaluations contrary to reality. On the other hand, it poses the question: is this finding contrary to the assumption that a child's performance level can be raised through raising his confidence level? Note that the experimental groups were fairly homogeneous in their actual ability. A greater range of ability in the sample would have undoubtedly resulted in a significant difference on the *WISC,* between groups labeled high and low in confidence. This does not obviate the fact that at certain levels of ability children may be expected to perform about the same, yet differ in confidence regarding their ability.

The median rhos of Table 1 reflect a more stable task interest value for the low confidence group. As expected, there is a halo-effect between the self-estimated ability ranking and the final ranking of task interest. Although not reaching significance, there appears to be a greater halo-effect for the low confidence group. This halo phenomenon more than likely bears a relation to the findings of Cartwright (2). It demonstrates that task interest changes as a result, not only of past success and failure of a definable sort (i.e., when *S* is told his score), but also as a result of past success and failure of a subjectively perceived nature. Neither experimental group estimated its actual ability with any proficiency. In a situation such as this, with little objective feedback or normative criteria to respond to, it is not surprising to find that *S*s cannot rank their abilities with any accuracy. Finally, a very significant chi-square is in line with the major hypothesis that the influence of task interest on self-estimated ability is greater in low than high confidence groups.

E. SUMMARY AND CONCLUSIONS

This study examines the influence of self-confidence on the relationship between self-estimated ability and task interest value. Other aims included the accuracy of self-estimated ability, the relationship of confidence level to actual ability, and the change in task interest as a result of experience. Subjects were 60 boys from an elementary school, and the tasks performed were four subtests of the Wechsler Intelligence Scale for Children.

1. The influence of interest value on self-estimated ability was greater in low than high confidence groups.
2. Both high and low confidence *Ss* performed at about the same level, according to their pro-rated *IQ*'s.
3. Neither experimental group estimated actual ability with any proficiency.
4. In general, the low confidence *Ss* demonstrated a more stable task interest.

REFERENCES

1. ANDERSON, ROSE G. Do aptitudes support interests? *Guid. J.*, 1953, **32**, 14-17.
2. CARTWRIGHT, D. The effect of interruption, completion, and failure upon the attractiveness of activities. *J. Exp. Psychol.*, 1942, **31**, 1-16.
3. DUNLAP, J. W. The predictive value of interest test items for achievement in various school subjects. *J. Appl. Psychol.*, 1935, **19**, 53-58.
4. FRANDSEN, A. Interests and general educational development. *J. Appl. Psychol.*, 1947, **31**, 51-66.
5. FRANDSEN, A., & SESSIONS, A. D. Interests and school achievement. *Educ. & Psychol. Meas.*, 1953, **13**, 94-101.
6. GEBHARD, M. E. The effect of success and failure upon the attractiveness of activities as a function of experience, expectation, and need. *J. Exp. Psychol.*, 1948, **38**, 371-388.
7. HIRSCH, N. D. M. Relationship between interest, ability, and self-estimated ability among maladjusted boys. *J. Abn. & Soc. Psychol.*, 1939, **34**, 395-399.
8. MCNEMAR, Q. Psychological Statistics. New York: Wiley, 1949.
9. THORNDIKE, E. L. Interests and abilities. *J. Appl. Psychol.*, 1944, **28**, 43-52.
10. TOWNSEND, A. Achievement and interest ratings for independent-school boys. *Educ. Rec. Bull.*, 1945, No. 43, 49-54.
11. WECHSLER, D. Manual: Wechsler Intelligence Scale for Children. New York: Psych. Corp., 1949.

Department of Psychology
Washington State College
Pullman, Washington

The Journal of Genetic Psychology, 1959, **94**, 161-168.

COMPARING THE EFFECTIVENESS OF THE EGO-STRENGTH *Q*-SORT TEST BY USE OF *R*- AND *Q*-METHODOLOGIES*

Phoenix College System, Phoenix, Arizona

RUSSELL N. CASSEL

A. PROBLEM

This study is concerned with comparing the effectiveness of The Ego-strength *Q*-sort Test for discerning between groups of individuals who have been hypothesized to have strong and weak ego-strength, in five separate situations, when scored by *R*- and *Q*-methodologies.

B. TEST AND CONCEPT

1. *The Ego-strength Q-sort Test*

A multiple revised edition of an experimental *Q*-sort test designed primarily to assess the ego-strength of typical individuals (1) is used. It is made up of 60 items which are to be rated by the subject taking the test on an 11-point scale that is normalized; so that only one item receives a rating of 1 or 11, three items receive a rating of 2 or 10, five items a rating of 3 or 9, seven a rating of 4 or 8, nine a rating of 5 or 7, and 10 a rating of 6.

All of the items on the test are positively stated and are presumed to be indicative of a happy person who has considerable ego-strength. This is done to accomplish two objectives: (*a*) to insure full use of the 11 points of the scale for all items (when items are stated both positively and negatively, research findings indicate that only the top and bottom ends of the scale are used for the different items, respectively), and (*b*) to provide a forced-choice test structure that will tend to minimize fakability of test scores by the subject (connotations of the various items do not provide a valid basis for the speculation of ratings which might yield vantage scores).

Items for the test were derived from a pool of 4,275 critical characteristic statements, solicited from several hundred graduate students who were majoring in counseling or guidance and in clinical psychology over a period of seven years, as being indicative of persons who were considered to be ideal examples of characteristically happy or depressed individuals. Statements

---

* Received in the Editorial Office on January 10, 1958.

in the pool were assigned to five different related groupings, and the items for this test were taken from each of the five groups; so that the number of test items were roughly proportional to the total number of items in the respective categories, and so that there was a representative coverage of the various sub-groupings within the five categories.

A listing of the test items by related groupings is given in Table 1.

TABLE 1

Category One—Ego-Status (self-sufficiency)

1. (15) Has feeling of being secure.
2. (17) Has general approval of self and an absence of shame.
3. (29) Seeks and accepts personal responsibility.
4. (35) Has high level of personal confidence or self-assurance.
5. (40) Accepts disappointments in own stride.
6. (51) Does not feel inferior to others.
7. (52) Is generally very careful of personal appearance.
8. (54) Has a very accurate appraisal of own ability.
9. (59) Is not afraid of own personal deficiencies.
10. (60) Does not boast of own accomplishments.

Category Two—Social Status (personal prestige)

1. (5) Gets personal satisfaction out of helping others.
2. (10) Has several close and intimate friends.
3. (11) Likes to do things with others.
4. (12) Has a feeling of importance among peers (friends).
5. (16) Seldom breaks the laws or rules of the group.
6. (21) Gets along well with others.
7. (27) Has systematically organized social activity (regular).
8. (28) Is able to put self in the place of others.
9. (33) Generally well accepted by members of own group.
10. (37) Is willing to accept appropriate authority.
11. (38) Is not usually suspicious of what others say or do.
12. (39) Is not envious of other person's accomplishment.
13. (43) Is trusted by others.
14. (49) Does not try to avoid others.
15. (53) Does not feel ill-at-ease with peers (friends).
16. (57) Is able to sense the disapproval of others.

Category Three—Goal Setting and Striving (ego-involvement)

1. (1) Actively engaged in setting personal goals.
2. (2) Continued recognized progress towards personal goals.
3. (3) A strong personal desire to achieve own goals.
4. (4) Has clearly designed goals for future endeavor.
5. (9) Personal goals are in line with own ability.
6. (19) Makes own personal decisions.
7. (24) Has high interest in own vocational pursuit.
8. (31) Continuously striving to reach personal goals.
9. (41) Tries to make a success in life.
10. (45) Often experiences success.
11. (46) Is conscientious, responsible, and painstaking in efforts.
12. (55) Believes life is important and purposeful.
13. (58) Has broad outlook and well rounded knowledge.

TABLE 1 (*continued*)

Category Four—Good Mental Health (satisfactory personal adjustment)

1. (6) Maintains a cheerful disposition.
2. (7) Has an active hobby in which he or she is interested.
3. (8) Adjusts to progressive change readily.
4. (14) Has developed an active outlet for creativity (self-expression).
5. (18) Has a satisfying and adequate home life.
6. (20) Is usually tolerant and not very sensitive or easily upset.
7. (22) Often receives praise or coveted honors.
8. (23) Is usually free of worry and emotional problems.
9. (25) Knows how to relax and practices it.
10. (26) Tends to see the good in things.
11. (30) Remains calm under pressure.
12. (34) A reasonably well adjusted person.
13. (36) Has a good sense of humor.
14. (42) Devotes a portion of own time to religion.
15. (44) Gets a fair share of the luxuries of life relative to group.
16. (47) Has an occasional exposure to adventure, risk, and danger.
17. (48) Has satisfactory psycho-sexual relationships (boy-girl problems).
18. (50) Places appropriate values on money and material things.
19. (56) Can express self well.

Category Five—Good Physical Status (good physical health)

1. (13) Has basic necessities of life.
2. (32) A reasonable degree of good physical health.

2. *Defining the Ego*

In general the ego is defined as that aspect of the personality which is used by the individual in maintaining contact with the outside world; it is presumed not to be present at birth, but it emerges gradually as the person grows older. It is fashioned and molded by the social climate in which the group membership is bathed; so that by the age of from 2 to 5 years old a working ego is present. The principal function of the ego is the implementation of activity designed to generate happiness and satisfaction for the individual.

3. *Construct Validity of Ego Concept*

A review of the professional literature relative to ego psychology was made by Bernard Hartman at Our Lady of the Lake College, San Antonio, Texas, recently as a portion of a thesis, and on the basis of this review he constructed an ego strength rating scale (2). The items used for rating purposes were: ego-status, social status, goal-setting, goal-striving, aspirations comparable to ability, psycho-sexual and emotional adjustment, religious and moral satisfaction, good sense of reality, satisfactory physical health, and happiness and contentment.

In the same thesis Hartman administered 10 different tests of ego strength and related assessment to a group of 52 Army Neuro-psychiatric Technicians, and then factor analyzed the data. Eight separate centroid factor loadings were extracted which were defined as follows: irreality, emotional status, social status, ego-status, mental health, goal setting and striving, psychosexual and affiliation, and intelligence.

The five parts of the present test indicate excellent agreement with the findings indicated in Hartman's study, and each of the separate parts appeared directly related to one of Hartman's ratings or extracted factors, respectively. There were items in the present test relative to all ratings and factors, in addition.

C. Groups Utilized

1. *Typical Personality Groups*

a. Typical young adults. One hundred individuals (about equally divided for sex) ranging in age from 17 to 31 years old with an *M* of 23.8 and with an *SD* of 3.79. Their last school grades successfully completed range from 10 to 19 with an *M* of 11.35 and with an *SD* of 1.8. The *IQ*'s range from 90 to 135 with an *M* of 103.7 and with an *SD* of 13.9. This group represented young adults available to graduate students at Our Lady of the Lake College in south central Texas.

b. Select male youth. One hundred Air Force Pre-flight Cadets, all of whom were carefully selected for purpose of commissioning and training as officer pilots. They ranged in age from 17 to 26 years old with an *M* of 22.4 and with an *SD* of 1.79. Their last school grade ranged from 12 to 19 with an *M* of 13.20 and with an *SD* of 0.92. Their *IQ*'s ranged from 115 to 150 with an *M* of 120.3 and with an *SD* of 8.91. They were all undergoing their preflight training program.

c. Air Force basic airmen. One hundred Basic Airmen representing portions of two available flights with no selective factors operating. Their ages range from 17 to 27 with *M* of 18.9 and with *SD* of 1.3. Their last school grade ranges from 9 to 14 with *M* of 10.5 and with *SD* of 2.4. Their *IQ*'s range from 88 to 141 with *M* of 97.0 and with *SD* of 15.6

d. Air Force basic women (WAF). One hundred Basic *WAF* persons from two available flights, and they ranged in age from 18 to 32 with *M* age of 19.9 and with *SD* of 3.2. Their last school grade ranges from 11 to 15 with *M* of 12.5 and with *SD* of 2.1. Their *IQ*'s ranged from 88 to 141 with *M* of 109.3 and with *SD* of 12.3.

e. Ordained religious ministers. Thirty ordained religious ministers with an average age of about 31 years; all with *IQ*'s over 115, all with five or more years of college, and with five or more years as an ordained minister in their respective faith. About half of the group were Catholic Priests, and the other half came from various other religions, including two Jewish Rabbis.

2. *Deviant Personality Groups*

a. Neuro-psychiatric hospital patients. One hundred patients in two general hospitals all diagnosed as schizophrenia (mostly simple stage and ambulatory), and about equally divided between the sexes. They ranged in age from 14 to 60 years old with *M* of 37.0 and with *SD* of 4.7. All of the subjects had a minimum of 5th grade education or better, and were estimated to have low average or better intelligence. The subjects were taken from a roster of simple schizophrenia patients, and those subjects who were unable to complete the test were eliminated from the study.

b. Federal reformatory prisoners. One hundred male adult prisoners currently committed to a Federal Reformatory in South Central United States, and all of whom volunteered to take some experimental personality and leadership tests. They ranged in age from 17 to 27 with *M* of 22.3 and with *SD* of 3.7. Their last school grade completely ranged from 6 to 16 with *M* of 10. and with *SD* of 2.7. Their *M IQ* was 92.0 with *SD* of 14.2.

D. COMPARATIVE PROCEDURE

Five different comparative assessments were made in this study with each different one involving a high and low hypothesized ego-strength group, respectively.

1. *R-Methodology*

Under this approach The Ego-strength *Q*-sort Test was scored by taking the sum of the ratings assigned to each of the five categories on the test (part scores) and then by computing an inter-correlation matrix using the Pearson technique. Then a point bi-serial r was computed for each of the part scores and a multiple point bi-serial correlation was developed through the use of a multiple discriminant function as described by Horst and Smith (3). The five indexes for the groups involved in this study are shown in Table 2.

2. *Q-Methodology*

Here the *Q*-sort array (ratings assigned to the 60 items on the *ESQT*) were correlated against the norms compiled in the test standardization (consisting of the mean *Q*-sort arrays for widely diversified groups of individuals, ranked and forced into a normalized distribution). The Pearson r obtained

TABLE 2
COMPARING SCORES ON THE EGO STRENGTH *Q*-SORT TEST ON THE BASIS OF *R*- AND *Q*-METHODOLOGY BETWEEN GROUPS OF HIGH AND LOW HYPOTHESIZED EGO STRENGTH

Groups	*Q*-Methodology†	*R*-Methodology††	Differences	Significance (t values)
Typical and deviant personality groups:				
1. Typical young adults (*N*–100) Vs. *N-P* patients (schizophrenias) (*N*–100)	.605**	.436**	.169	3.84**
2. Select male youth (*N*–100) Vs. Federal Reformatory prisoners (*N*–100)	.392**	.341**	.051	0.77
Typical personality groups:				
3. Ordained religious ministers (*N*–30) Vs. Air Force basic airmen (*N*–100)	.412**	.625**	.213	3.05**
4. Air Force basic *WAF* (*N*–100) Vs. Air Force basic airmen (*N*–100)	.396**	.406**	.010	0.14
Deviant personality groups:				
5. Federal Reformatory prisoners (*N*–100) Vs. *N-P* patients (schizophrenias) (*N*–100)	.221**	.382**	.161	2.06*

† For *Q*-Methodology indexes are point bi-serial r coefficients.
†† For *R*-Methodology indexes are multiple point bi-serial R coefficients.
* Statistically significant at 5 per cent level.
** 1 per cent level.

was then converted to a Fisher z' score (standard score), and a point bi-serial r was computed for each of the five separate groups; the resultant coefficient was then reconverted to an equivalent r. These indexes are illustrated in Table 1 for all five groups.

E. FINDINGS

1. *Comparisons Between Typical and Deviant Personality Groups*

By use of both *R*- and *Q*-methodologies the *ESQT* discerned reliably at the 1 per cent level of confidence between groups of individuals with deviant and typical personality characteristics. This reliable discerning was present when the typical young adults were compared with *N-P* subjects, and when select young adults were compared with Federal Prisoners in a Reformatory. In both groups the use of *Q*-methodology discerned more reliably, but only with the *N-P* subjects and typical young adults was the difference significant, with a t of 3.84 (reference Table 1).

2. *Comparisons Between Typical Personality Groups*

Here again both the *R*- and *Q*-methodologies discerned reliably at the 1 per cent level of confidence between low and high hypothesized ego-strength groups. In both typical group investigations (Religious Ministers & Airmen, and *WAF* & Airmen) the *R*-methodology discerned more reliably, but only with the Religious Ministers & Airmen was the difference statistically significant at the 1 per cent level of confidence.

3. *Comparisons Between Deviant Personality Groups*

Here again both the *R*- and *Q*-methodologies on the *ESQT* discerned reliably at the 1 per cent level of confidence between *N-P* subjects and Federal Prisoners. Here the *R*-methodology was more effective with a difference significant at the 5 per cent level of confidence (t equals 2.06).

F. Conclusions

Findings from this study indicate that the *ESQT* discerns reliably between groups of individuals with hypothesized strong and weak ego-strength, regardless of whether or not one of the groups is classified as a deviant personality one. The *Q*-methodology tends to favor deviant personality groups; while the *R*-methodology tends to favor the typical personality ones. Where both groups are from the deviant personality area, the *R*-methodology appears favorable in discerning.

G. Summary

This study was concerned with describing a newly developed ego strength test based on *Q*-methodology in which 60 items were taken from a pool of 4,275 critical characteristics indicative of characteristically happy persons. The items came from five related categories of such statements which constitute the various part scores on the test, i.e., ego-status, social status, goal-setting, mental health, and physical status. These five parts to the test correspond rather closely to an extensive review of ego psychology literature, and with a factor analysis study of 10 ego-strength and related tests. Scoring on the test was accomplished by both R- and Q-methodologies. In the first a sum of the ratings assigned to the items within the various categories was utilized, while in the latter a correlation of the Q-sort array was made with the norm compiled in the standardization of the test.

Comparative assessments were made in five different situations and in each one a group of hypothesized strong ego-strength individuals were compared with a similar group of hypothesized weak ego-strength ones. Both deviant (*N-P* subjects and Federal Prisoners) and typical personality groups were involved. In all instances scores from the *ESQT* discerned between the two groups reliably at the 1 per cent level of confidence. For the deviant personality comparisons with typical personality groups, the Q-methodology was more favorable; while with the typical personality, or where two groups both from deviant personality groups were involved, the R-methodology was more favorable. Differences between the R- and Q-methodology were not always significant in referenced directions.

REFERENCES

1. CASSEL, R. N. The Ego Strength *Q*-Sort Test and Manual. Chicago 1, Illinois, Psychometric Affiliates, 221 N. LaSalle St., 1958.
2. HARTMAN, B. The Ego Strength Rating Scale. San Antonio, Texas: Our Lady of the Lake College, 1957.
3. HORST, P., & SMITH, S. The discrimination of two racial samples. *Psychometrika,* 1950, **15**, No. 3.
4. MOWRER, O. H. *Q*-Technique: Description, history, and critique. *In* (O. H. Mowrer, *Ed.*) *Psychotherapy Theory and Research.* New York: Ronald Press, 1953.

4532 North 53rd Drive
Glendale, Arizona

The Journal of Genetic Psychology, 1959, **94**, 169-179.

A DEVELOPMENTAL ANALYSIS OF PROVERB INTERPRETATIONS*

Department of Child Study, Vassar College

Claudia Richardson and Joseph Church

A. Problem

The task of defining common proverbs has been used to good advantage in the study of thinking (6), for psychometric purposes (4, 5, 8, 9), and for clinical diagnosis (1, 7, 10). Previous investigations, however, have not been altogether satisfactory for students of the development of thought. In some studies (5, 6), data were gathered by means of a multiple-choice technique, which imposes certain obvious limits on the possible range of responses. In other studies, although the subject was free to respond as he wished, scoring has been restricted to only one or two aspects of a response: correctness (8), or literal vs metaphorical interpretation (5, 9), which Gorham (4, 5) designates concrete vs abstract. In clinical applications, attention is likely to be directed to "bizarre" or "projective" features of responses, with little reference to standards of normal behavior (1, 7, 10).

The aim of the present study is both methodological and substantive. Methodologically, it seeks to arrive at a set of scoring categories which (*a*) are reliable, (*b*) take account of qualitative features of responses, (*c*) permit statistical comparisons, and (*d*) more nearly exhaust the information contained in responses than do earlier scoring systems. Substantively, it seeks to shed some light on certain shifts in cognitive functioning that take place in the course of development. It should be stressed that the method of age-group comparisons used in this study is not aimed at defining age norms, although such norms could conceivably be established by applying the categories of this study to a larger, more representative standardization sample. The goal, rather, is the specification of relatively mature and immature response characteristics which, though in some ways peculiar to the present task, are in other ways relevant to all of cognitive development. One may further hope, following the tenets of general developmental theory (11), that these findings will be applicable to abnormal as well as normal behavior.

B. Procedure

Sixty-four children, aged 7:10 to 12:5, were asked to give the meanings of seven proverbs. To provide further material for comparison, the same

---

* Received in the Editorial Office on January 21, 1958.

seven proverbs were presented to 30 normal "adults" between the ages of 14 and 67 years. The children were drawn from four sources representing a wide spread in geographical and socio-economic backgrounds: a large urban public school in Denver; a small private, progressive school in Poughkeepsie, N. Y.; a public school in the semi-rural community of Hyde Park, N. Y.; and a community recreation center in a relatively depressed neighborhood in Poughkeepsie. Unfortunately, it was impossible to obtain measures of the children's intellectual standing, but we have no reason to suppose that they deviate markedly, as a group, from the normal range. The children were assigned to four groups, each comprising an age span of one year (ages 8, 9, 10, 11), except that two seven-year-olds were added to the youngest group and two 12-year-olds to the oldest. One subject had to be eliminated from the youngest group because his monosyllabic responses defied interpretation. The distribution of the groups by age, sex, and source is shown in Table 1.

TABLE 1
DISTRIBUTION OF SUBJECTS BY AGE, SEX, AND SOURCE

	N	Age range	Mean age	Boys	Girls
Group I	16	7:4-8:7	8:4	7	9
Group II	15	9:0-9:11	9:5	7	8
Group III	15	10:0-10-11	10:5	9	6
Group IV	17	11:0-12:5	11:6	12	5

	Denver Public School	Pkpsie Private School	Hyde Pk Public School	Pkpsie Rec. Ctr.
Group I	3	5	4	4
Group II	10	0	3	2
Group III	6	8	0	1
Group IV	5	4	6	2

The seven proverbs used were: An ounce of prevention is worth a pound of cure; Don't cross your bridges until you get to them; Every cloud has a silver lining; All that glitters is not gold; Don't cry over spilt milk; Where there's smoke there's fire; and You can't teach an old dog new tricks. (For brevity's sake, these proverbs will hereafter be referred to, respectively, as Ounce, Bridges, Cloud, Milk, Smoke, and Dog.) The specific criteria observed in selecting the present set were that they be in widespread use and that they be phrased figuratively rather than literally.

The proverbs were read aloud by *E* (Richardson) to individual *S*s. Responses were recorded verbatim in shorthand. After *E* had enlisted a child's coöperation and moved with him away from the group, the following instructions were given:

> I have a list of proverbs here. Do you know what a proverb is? Well, it's an old, old saying. What I want you to do is tell me what you think each saying means. I'm going to write down what you say so I can remember it. There's no right or wrong answer—all I want to know is what you think they mean.

After each response, *E* commented "Good" or "Fine—and now the next one." The proverb was repeated as often as necessary. If *S* asked, the word "prevention" was defined; otherwise no help was given beyond encouragement to guess rather than give up. Those children who asked about the purpose of the experiment were told that different people thought the proverbs meant different things, and that *E* was interested in collecting different answers; this reply was apparently quite satisfactory. In general, *S*s' interest and coöperation were high.

C. Scheme of Analysis of Data

Three sorts of information may be gleaned from the responses obtained: (*a*) how *S* understood the task (as opposed to how he understood the proverb); (*b*) what *S* understood the proverb to mean; (*c*) the means at his disposal for communicating his understanding. Although the material we have to work with consists entirely of the end-product of (*c*), it is (*a*) and (*b*) which lend themselves most readily to analysis and with which this report is concerned.

1. *Task Orientation*

The way in which *S*s understood the demands of the task, rather than the meaning they attached to the proverbs, is termed Task Orientation. Many proverb interpretations become comprehensible only when we realize that *S* is answering a question he has not been asked. For statistical purposes, there are two possible scorings under Task Orientation: Definition, indicating that the response is designed to say what the proverb means, and Other.

A considerable variety of task orientations is found under the heading of Other. Among the questions that *S*s attempted to answer are: (*a*) Is this proverb true (or valid)? E.g., Ounce, F9:0, "I think it is"; answers consisting of an unelaborated Yes or No are classified thus. (*b*) Why (or how do we know, or under what conditions) is this proverb true (or valid)? E.g., Cloud, F7:11, "Because it's made of thin air?"; Gold, *id.,* "Because copper looks like gold"; Ounce, M8:3, "Cuz medicine will keep you well." Typically, but not always, such responses begin with "because"; it must be noted, too, that a number of these responses set out to dispute the proverb as well as to justify it. (*c*) How can we explain the situation represented by the proverb? E.g., Cloud, M8:5, "Well, that's when the sun reflects on it and makes a ring around it." (*d*) What does one do in the situation depicted by the proverb? E.g., Milk, M8:6, "Just wipe it up"; Dog, *id.*, "Do you teach them the same tricks over and over?" (*e*) What function does the proverb serve? E.g., Smoke, F7:11, "So you know not to burn yourself." In addition to these modes of task orientation, some *S*s saw the proverb as a riddle to be solved, others as a test of information (e.g., How many ounces in a pound?); a few *S*s treated the task as a word association test.

2. *Proverb Comprehension*

In the matter of how *S* understood the proverb, no account is taken of the correctness of responses. In point of fact, very few child *S*s gave "correct" interpretations. Proverb comprehension is scored along three dimensions: Specific-General, Literal-Figurative, and Physiognomic-Articulated. A dimension is here defined as a set of categories ranged between two polar extremes. It may, of course, consist only of the two extremes, or, as in the case of physiognomic responses, simply of the presence or absence of a given response characteristic. It does not imply a continuum. Because, empirically, these dimensions represent developmental polarities as well as logical ones, they are not orthogonal but parallel. They are, however, independent, in that the position assigned a response on one dimension in no way determines its position on another.

a. Specific-General Dimension. This dimension includes two categories, General Statement and Concrete Specification. The first is self-explanatory, since virtually all proverbs can be reduced to the form, "Given a situation of Type *A*, then event *A'* ensues"; i.e., they are general statements of antecedent-consequent relations. Several degrees of generality of interpretation are, of course, possible (it has proved quite simple to score responses as under-specific or over-specific), but any response which refers to a class of events, no matter how narrow, is scored as General Statement.

Concrete Specification responses, on the other hand, are those in which a proverb is interpreted as describing a particular person, thing, attribute, situation, or event, almost as though the proverb were a picture caption or a story title or synopsis. E.g., Smoke, F7:6, "Fire in a fireplace"; Dog, F8:5, "Well, there's an old dog, and you're trying to get him to sit up or something, and he won't sit up for you"; Cloud, M8:6, "It means that the cloud's dark"; Cloud, M9:6, "Maybe it's a man that has a silver lining in his coat." Two minor scoring difficulties arise in connection with this category. First, a tendency to concrete specification is apparent in such responses as Dog, F9:5, "It's hard for the dog . . . cuz he's so slow and old"; in keeping with a conservative scoring policy, such responses are classed as general statements. Second, concrete specifications must be distinguished from responses consisting entirely of examples, e.g., Gold, M11:4 "Silver. A mirror in the sun reflecting, oh boy, a lot of things—diamonds, tinfoil."

b. Literal-Figurative Dimension. Three categories are needed to account for *S*s' differing sensitivity to the metaphorical phrasing of the proverbs: Literal, Mixed Literal and Metaphorical, and Metaphorical. In a literal interpretation, *S* merely takes the words of the proverb at their face value. Mixed interpretations are those in which the proverb is given a metaphorical meaning but certain literal elements remain, as in Gold, M10:5, "Everything you see that is shiny doesn't mean it's worth a million dollars" (over and above the literalism of *shiny* in this example, any reference to monetary value in defining Gold, or to the practice of medicine or the measurement of weight in Ounce, is considered literal; however, when *S* gives the literal proverb as one example of a more general relationship, this is not treated as literalism); Cloud, F10:6, "Might be something in that cloud that would be valuable." This category is also a receptacle for responses having so low a degree of specificity that their position on this dimension is wholly indeterminate, e.g., Bridges, M10:4, "If you are doing something, don't do anything before you get to this thing." Metaphorical interpretations are those in which the literal meanings of the proverbs' key words do not appear. The metaphorical category could be subdivided to produce a fourth category, Analogical Metaphor, containing those responses in which the relationships among terms in the original proverb are preserved in the metaphorical interpretation; this category is chiefly useful, however, only in distinguishing between the metaphorical interpretations of children and adults.

c. Physiognomic-Articulated Dimension. This dimension comprises the two polar categories Physiognomic and Articulated. The latter is scored by exclusion. Any one of three criteria is used in scoring responses as physiognomic: (*a*) The proverb meaning is dominated by a single term, to the neglect of the remainder of the proverb. E.g., Cloud, M7:4, "There's a blue cloud in the sky." (*b*) The proverb meaning is dominated by the "atmosphere" of the proverb, with little or no regard to any of its terms. E.g., Ounce, M9:2, "Well, if something is too heavy for an old person, you could help him carry it" (cf. Ounce, M9:4, "If a board was too heavy for my father, I would help him carry it"); Cloud, F11:6, "Someone's being nice to everybody else, not being piggy." (*c*) A concept clearly alien to the terms of the proverb appears in the definition. E.g., Bridges, M9:6, "Don't go on a bridge that's been torn down."

D. Reliability of Categories

All responses were scored independently by the two authors. Of 1656 scorings (414 responses times four scoring dimensions), there was independent agreement on 1615, or slightly more than 97 per cent. Of the 41 disagreements, 30 were easily resolved by discussion, and 6 proved unscorable. The authors continued to disagree on five scorings; in these cases, responses by *S*s in the two younger groups were assigned to the more mature category, and those by older *S*s to the less mature.

E. Statistical Methods

Comparisons between groups were made by means of the *t* test, following the conservative method suggested by Edwards(3), following Snedecor, and Cochran and Cox, for use when the *n*s and the variances of the groups differ. It should be noted that developmental studies often contain an implicit hypothesis that group variances will differ, since the youngest *S*s often display homogeneously immature processes, the oldest more nearly homogeneously mature ones, and intermediate *S*s a mixture of mature and immature processes. This method of calculating *t* seems, in the present instance, roughly to double the probability values given by the usual method. A further note is in order: all scoring dimensions contain one more degree of freedom than is immediately visible, consisting of the no-response category; for this reason, polar category pairs may not strictly mirror each other.

F. Findings

As shown in Tables 2–5, there are significant changes with age in the frequency of response along all dimensions except Physiognomic-Articulated. Table 2 shows a steady rise in orientation towards the task as one of definition, with a significant increase in Definition responses between Groups III

TABLE 2
Mean Frequency of Definition and "Other" Responses

	Definition	t	"Other"	t^{a}
Group I	1.69		4.62	
		1.3		.87
Group II	2.60		4.00	
		1.8		1.98
Group III	3.93		2.53	
		2.2*		1.68
Group IV	5.41		1.41	

[a] Group I vs Group III, $t = 2.9^{**}$; Group II vs Group IV, $t = 3.86^{**}$.
* Probability less than .05.
** Probability less than .01.

and IV. There is a corresponding decrease of "Other" responses; while there are no significant differences in this category between adjacent groups, the over-all shift is highly reliable. There is a significant decline with age (Table 3) in Concrete Specifications (the reversal in Group IV is the work of a single *S* and does not alter the general trend), with a sharp decline occur-

TABLE 3
MEAN FREQUENCY OF CONCRETE SPECIFICATIONS AND GENERAL STATEMENTS

	Concrete specifications	t[a]	General statements	t
Group I	1.50		4.87	
				.8
Group II	1.33		5.23	
		2.71*		3.2*
Group III	.47		6.07	
				.15
Group IV	.82		6.00	

[a] Combined Groups I & II vs combined Groups III & IV.
* Probability less than .01.

TABLE 4
MEAN FREQUENCY OF LITERAL, MIXED, AND METAPHORICAL RESPONSES

	Literal	t	Mixed	t	Metaph.	t
Group I	6.12		.06		.06	
		1.68				
Group II	5.23		.53		.47	
		2.61*		2.1*[a]		4.2**[a]
Group III	3.27		.80		2.40	
		1.93		7.7**		1.3
Group IV	1.82		1.65		3.35	

[a] Combined Groups I & II vs Group III.
* Probability less than .05.
** Probability less than .01.

TABLE 5
MEAN FREQUENCY OF PHYSIOGNOMIC RESPONSES

	Total physiog.	Literal physiog.	t	Metaph. physiog.	t
Group I	1.50	1.38		.06	
					4.1*[a]
Group II	2.20	1.20		1.00	
			3.5*[b]		
Group III	1.73	.53		1.20	
					2.8*[c]
Group IV	2.53	.24		2.29	

[a] Group I vs combined Groups II & III.
[b] Combined Groups I & II vs combined Groups III & IV.
[c] Combined Groups II & III vs Group IV.
* Probability less than .01.

ring between Groups II and III; there is a corresponding rise in the frequency of General Statements. Literal interpretations (Table 4) decrease with age, with the greatest change coming between Groups II and III. There is an increase with age in the frequency of Mixed interpretations, these being all but non-existent in the youngest group. There is an even more marked increase with age in Metaphorical interpretations, with the most pronounced change between Groups II and III.

Table 5 shows not only the mean frequency of physiognomic responses, which contrary to expectations, tend to increase with age, but also a further breakdown of physiognomic responses classified according to whether they are literal or metaphorical in content. Thus divided, the data show that there is a significant decline with age in the frequency of physiognomic literal responses, and a significant increase in physiognomic metaphorical responses. In short, a major portion of metaphorical interpretations are of a physiognomic kind.

1. *Adult Interpretations*

Analysis of responses given by the sample of "adults" indicates a continuation of the trends revealed in children's responses, except, as suggested earlier, the adult metaphorical interpretations are less physiognomic and better preserve an analogical relationship with the proverb than do those ot children. It should be noted, however, that all the immature forms of response found in children also occur, although with decreased frequency, in adults. That is, while new forms of response emerge with increasing maturity, earlier forms still persist. In addition, many adults begin with a rather primitive interpretation and then progress microgenetically to a more mature one.

G. Discussion

Since the present study is empirical, its implications can be only tentative. However, the categories used here have been applied to roughly comparable data gathered in another context by Church (2), and the results are in good agreement. Perhaps the most important methodological implication of the present research is that the correctness of an *S*'s definitions can be totally disregarded in an analysis of his performance.

The data suggest several matters of substantive interest. In terms of task orientation, it is evident that the younger *S*s do not maintain a stable problem schema but reinterpret the task from moment to moment. This fact seems related to two immature tendencies. First, there is the diffuseness and lability of the child's thinking. Second, it appears that the notion of definition is alien to the young child, for whom experience is "real" (in Piaget's sense) and obvious; it is less something to be analyzed and explained than to be taken for granted and reacted to. Although, by age eight, children have lost much of their realism with regard to the material world, they still retain it when it comes to dealing with language—they are capable of a metaphysics but not of a metalinguistics. In general, it seems that the younger *S*s gave definitions only when they themselves were baffled by a proverb; in other words, they gave definitions for their own benefit rather than because this was what the instructions demanded.

One interesting finding is that in certain respects the younger *S*s perform better than older ones, especially as regards susceptibility to physiognomic impressions. Their superiority seems related to the fact that they are dealing with simpler materials uncomplicated by a double layer of meaning.

Children's tardiness in perceiving proverbs as metaphorical, and their subsequent difficulty in establishing a stable analogical relationship between literal and figurative meaning, is in seeming contrast to Werner's (11, pp. 152-65) observation that metaphor is a basic, primitive, spontaneous mode of expression. Indeed, children much younger than those in the present sample speak in what sound like strophes. In effect, it appears that we are dealing with two different types of metaphor. The primitive metaphor seizes on gross phenomenal points of similarity between referent and figure, a similarity that may be quite peripheral to the essential nature of either; it is likely to appear in a context compounded of literal statement and figurative allusion; and it is often freighted with connotations that permit the movement of thought in many directions. Most important, it is not conceived by the speaker as metaphorical; as in a dream, several levels of meaning may be syncretically condensed in a single image. The mature mataphor, on the other hand, consists of a concrete analog of a general proposition, an analog not limited to the clearly figurative terms but including their relational context as well, making for a statement whose literal meaning, at least, is fairly unequivocal. Both types, of course, substitute concrete imagery for abstract, categorical language.

A recurring criticism of developmental studies of cognitive behavior is that the findings may represent the child's ignorance rather than a fundamentally different mode of organization and operation. It is undoubtedly true that many *S*s did not know the meanings of certain words (especially "prevention"). But the *S*s themselves were largely unaware of their ignorance —when they were aware of ignorance, it was ignorance of the proverb as a whole, and not of its lexical units. In only two cases did an *S* ask for clarification of a particular term, and in only seven responses was there evidence than an *S*'s difficulty had a focus in the misinterpretation of one word.

Future work in this area might concentrate initially on the stabilization and further elaboration of scoring dimensions and categories. It might be noted that other categories were suggested by the data of the present study but did not occur with sufficient frequency to justify inclusion. A different sampling of subjects or of proverbs might have brought some of these categories into greater prominence. It is quite evident that different proverbs favor the appearance of different types of response, although none of the proverbs used here was perfectly correlated, positively or negatively, with any one response type. Additional studies might deal with sex differences, intellectual differences, personality differences, performance of aged *S*s, and pathological manifestations.

H. Summary

Interpretations of seven common proverbs were obtained from 64 children, ages 7:4–12:5, and from 30 adults. Analysis of the responses revealed significant developmental changes in orientation to the task as one of definition rather than of evaluation, explanation, conundrum-solving, free association, etc.; in interpretation of the proverbs as general statements instead of concrete representations; and in comprehension of the metaphorical nature of the proverbs and ability to coordinate the two levels of meaning. With increasing appreciation of the complexity of the task, new obstacles to performance arise, so that certain deficiencies of performance are actually indicators of advanced maturity and the liabilities it brings. These findings were discussed relative to Werner's general developmental theory.

References

1. Benjamin, J. D. A method for distinguishing and evaluating formal thinking disorders in schizophrenia. *In* (J. S. Kasanin, *Ed.*) *Language and Thought in Schizophrenia.* Berkeley: Univ. California, 1944. (Pp. 65-90.)
2. Church, J. Children's resolutions of verbal contradictions: A developmental study. Unpublished doctoral dissertation, Clark Univ., 1954.
3. Edwards, A. L. Experimental Design in Psychological Research. New York: Rinehart, 1950. (Pp. 167-8.)
4. Gorham, D. R. A proverbs test for clinical and experimental use. *Psychol. Rep.*, 1956, **2**, 1-12, *Monog. Suppl.* 1.
5. ———. Clinical Manual for the Proverbs Test. Louisville: Psychol. Test Specialists, 1956.
6. Piaget, J. La pensée symbolique et la pensée de l'enfant. *Arch. Psychol., Genève.*, 1923, **18**, 273-304.
7. Rabin, A., & Broida, D. Projection via proverbs: Follow-up of a suggestion. *J. Consult. Psychol.*, 1948, **12**, 246-50.

8. TERMAN, L. M., & MERRILL, M. A. Measuring Intelligence. Boston: Houghton Mifflin, 1937. (Pp. 283-4.)

9. WECHSLER, D. Manual for the Wechsler Adult Intelligence Scale. New York: Psychol. Corp., 1955.

10. WELLS, F. L., & RUESCH, J. Mental Examiners' Handbook. New York: Psychol. Corp., 1945. (Pp. 115-7.)

11. WERNER, H. Comparative Psychology of Mental Development. New York: Int. Univ. Press, 1957.

Department of Child Study
Vassar College
Poughkeepsie, New York

The Journal of Genetic Psychology, 1959, **94**, 181-198.

ANXIETY REDUCTION AND NOVELTY AS GOALS OF VISUAL EXPLORATION BY MONKEYS*

Department of Physiology, Yale University School of Medicine

DAVID SYMMES[1, 2]

A. INTRODUCTION

Substantial evidence exists that monkeys as well as other species find reinforcement in exploration of the environment visually, aurally, or tactually. In other words, an empirical law of effect operates in connection with exploratory behavior and a number of discriminatory and maze habits can be learned with no other reward than the opportunity to explore (2, 4, 7, 14). Considerable controversy, however, has arisen concerning the detailed nature of the reinforcing events themselves and the conceptual steps necessary to arrive at an adequate theory of the source of the exploratory motive. The rôles of perceptual isolation (3, 6, 18), novelty in the current situation (1, 10, 15), and anxiety reduction (5, 16, 19) in producing or strengthening the exploration motive have been discussed in connection with various experimental results, but a number of unsettled questions remain.

The experiments described here were undertaken with the aim of illuminating three general questions about exploratory behavior in the monkey: (*a*) the motivational basis of exploratory behavior, especially the possible rôle of anxiety reduction; (*b*) the nature of the stimulus events capable of reinforcing preceding exploratory responses; (*c*) the development of a reliable quantitative measure of amount of exploratory behavior to standard incentive conditions. The last stems from our interest in developing indicators which may be useful in detecting minimal behavioral changes associated with stimulation and destruction of brain areas. Some evidence exists that cortical ablations diminish exploratory behavior in the monkey (9, 12).

The method chosen for the study of exploratory behavior in the monkey here involves measurement of spontaneous rate under controlled external stimulus conditions. This method is thought to maximize the opportunity for drive-arousing mechanisms associated with isolation to operate (minimum

---

* Received in the Editorial Office on January 27, 1958.

[1] The author acknowledges with gratitude the advice and criticism of Professor J. M. R. Delgado during the conduct of this research. Research facilities were provided in part by grants from The Foundations' Fund for Research in Psychiatry and the Office of Naval Research.

[2] Public Health Research Fellow of the National Institute of Mental Health.

interruption by *E*), and to diminish the chance of a single motivational source (anxiety) concealing the contribution of other less potent sources (desire for novelty). Two experiments are described which differ only in the type of response the monkey made to obtain visual access to a richer environment.

B. METHOD

1. *Subjects*

Eight experimentally naïve rhesus macaque monkeys were employed, ranging in weight from 1.9 to 4.8 kilograms. Four were used in Experiment I (Nos. 29, 32, 31, 74), and four in Experiment II (Nos. 85, 86, 87, 78). Two animals (Nos. 29 and 32) were tested first in Experiment I and then in Experiment II, making a total of 6 *S*s in the latter experiment. One *S* (No. 29) had been implanted with two multi-electrode needles 8 months previously and some stimulation of limbic structures had been carried out. No lasting behavioral effects were noted.

2. *Pretest*

The monkeys were housed singly in a colony of about 35 animals. No special preliminary training was undertaken beyond adapting *S*s to the carrying cage used to transport them to the testing room. All animals were fed a normal laboratory ration between 4 and 5 P.M. daily.

3. *Apparatus*

A test arrangement similar to that described by Butler and Alexander (8) was used. A wire mesh cage 30x30x25″ high with front and back of plywood was enclosed with wallboard to make a lightproof test box. A sliding panel at the rear provided a means of entrance and exit for *S*s. A drinking cup was placed on one wall and ventilation plus a masking noise was furnished by a duct leading from an air conditioner through the roof. Temperature was maintained between 70° and 80°F. inside the box with no consistent change during the test sessions. A single 6-watt fluorescent light in the roof at the rear illuminated the test box. A very small crack adjoining the light receptacle permitted *E* to observe *S*s occasionally during the test sessions. While the differential illumination inside and outside this aperture greatly favored *E* observing *S* over the reverse, all *S*s at some time during testing evidenced some interest in it, suggesting that they had visual or auditory cues differentiating it from other parts of the box.

For Experiment I a vertically hinged door was fitted to a 4x4″ window in the front wall of the box 11″ above the floor. The door was normally closed by light spring tension and a small permanent magnet. When pushed open from the inside the door permitted a leaf switch to close making a permanent record on an inkwriter operating on the other side of the room. A capacity relay connected across a large aluminum plate on the rear wall of the box and the wire mesh floor provided a record of activity via a second pen on the inkwriter. This counter operated so as to record one count whenever *S* passed within a few inches of the back of the box.

For Experiment II a hole 1″ in diameter was drilled in the front wall close to the door and at the same height. A baffle or shroud was fitted over the outside of the swinging door so that it could be opened as before but exposed only darkness. Visual access to the outside environment was possible only via the peephole. A light source, red filter, mirrors, and photoelectric relay were so arranged that when the monkey looked through the peephole a dull red beam was interrupted and the relay closed. The details of this mechanism are schematically illustrated in Figure 1. The relay was connected to the inkwriter and a permanent record taken as before. A con-

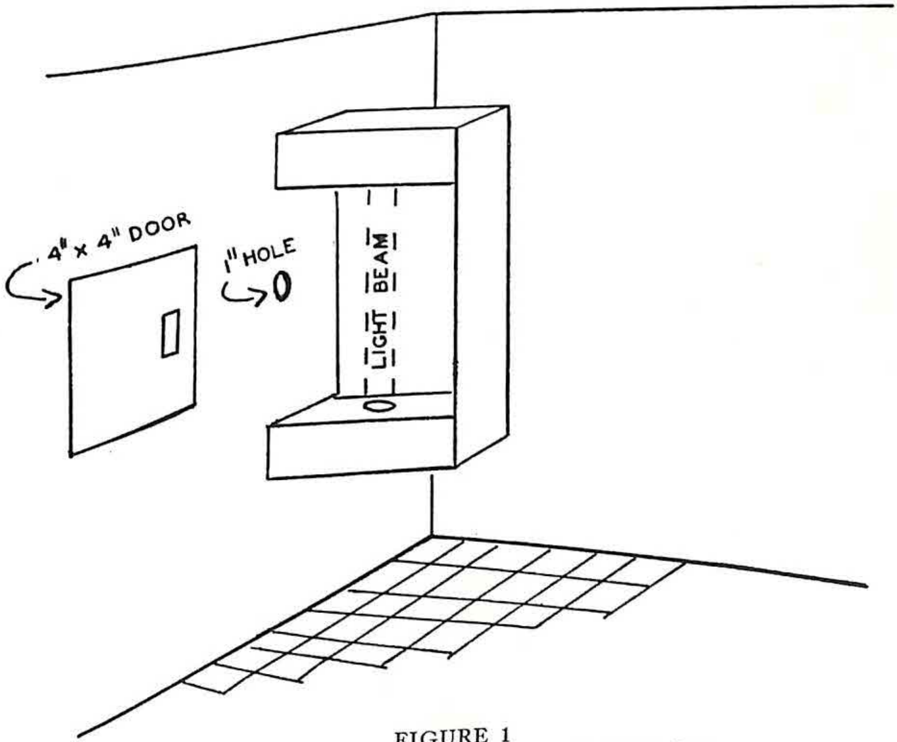

FIGURE 1
Schematic Drawing of Inside of Test Box

current record of responses on the door was taken, and the activity counter was disconnected altogether. The peephole and the door were so placed that *S* could, and often did, hold open the door and look through the hole simultaneously.

The visual field exposed by either the swinging door in Experiment I or the peephole in Experiment II consisted of a plain painted wall on which a white cardboard screen 27″ by 18″ was centered in front of the apertures at a distance of 48″. This screen was brilliantly illuminated by the field of a 35 mm. slide projector mounted on the roof of the test box. No other objects were visible from the door (which was partially barred with nails to prevent *S*s from putting their heads out) or the peephole, and the contrast of the screen against the dim background was extreme. The test room itself was illuminated only by the light reflected from the screen.

4. *Procedure*

Testing was carried out every other day beginning between 9 and 10 A.M. In Experiment I a period of adaptation and training to respond on the door was found necessary. All *S*s showed an initial reluctance to approach the partly open door and chose to hide at the rear of the box avoiding, if possible, the line of sight of *E* outside the box. To encourage them to use the door *E* was careful to stay out of sight and the door was left partly open with the spring disconnected, but no food or other special rewards were used. This period of preliminary training lasted between two and four days. When a criterion of several hours of more or less continuous door pushing activity under standard conditions was reached, formal testing began. In Experiment II no such adaptation period was required. All four naïve *S*s and the two experienced *S*s from Experiment I used the peephole within two minutes of their first exposure to the test box. Most ran over immediately and looked out the hole. Accordingly formal testing began in Experiment II from the first moment *S*s were in the test box. All *S*s also began to open the door after a variable period (usually within an hour) even though there was no visual reward.

Formal testing consisted of recording behavior during a 3½ to 4-hour test session for a total of 12 to 20 test sessions for the various *S*s. Each test session was broken down according to the following general plan: one hour of exposure to the standard stimulus condition (plain white screen); one hour during which one of seven altered stimulus conditions prevailed; and one final hour of standard condition. The list of altered stimulus conditions follows: (*a*) *Control,* in which the standard blank slide was momentarily removed and replaced; (*b*) *Darkness,* nearly total darkness outside the box, relieved only by dim light falling from the window when *S* opened his door; (*c*) *Triangle,* the image of an equilateral triangle 8″ on a side in solid black against the white background; (*d*) *Moving Triangle,* the same triangular image induced to move in an elliptical orbit filling the screen and completing a revolution every two seconds; (*e*) *Photograph,* a high quality color photograph of a monkey projected to approximately 3× life size. The photograph showed the monkey being held by a man's hand at the back of the neck and exhibiting a pronounced expression of fear; (*f*) *Live Monkey,* in which a small (1.6 Kg.), vocal, active monkey was brought in via the carrying cage and positioned 2 ft. from the test box window; (*g*) *Novelty,* a series of 40 slides containing a wide variety of outdoor scenes, histological material, *EEG* tracings, photographs of monkeys, cats, and birds, etc., projected automatically at the rate of one every 90 seconds. (All *S*s were adapted to the sound of the slide changing mechanism prior to testing.)

In detail, the following testing routine was employed: *S* was placed in the test box and behavior recorded for 60 minutes under standard stimulus conditions. Following this, and after a variable delay of 0 to 30 minutes (to discourage learning to temporal cues) a stimulus change was made, presumably with perceptible auditory cues for *S* in the box. The test period then began with the next response by *S* (the first occasion *S* had to see the change) and data were recorded as before. An exception to this procedure existed in the case of stimulus *live monkey,* when *S* was presumed to be aware from many cues that another monkey was outside and the test hour began immediately with the introduction of the object monkey. Under other conditions the duration of this *S*-controlled variable delay was usually less than one minute but occasionally as long as 10 minutes. At the end of the test hour the stimulus was returned to standard, *S* having had exactly 60 minutes from his first response in which to explore it. The final hour was again recorded from the first response of *S* to the standard stimulus. A variable period of 0–30 minutes concluded the test session, at which time *S* was removed from the test box and returned to his living cage.

In Experiment I, the first test session was always *control,* and then all the various stimuli were presented in random order, one per day, through the remaining test sessions. A total of five *control* days was obtained and three days for each experimental stimulus. Not all stimuli were presented to all *S*s, as will appear from the results section below. In Experiment II, the first four test sessions were always *control* (to equalize approximately degree of familiarity with the test box for *S*s in both experiments before any stimulus changes were introduced) and then the experimental stimuli were randomly presented as before for a total of three days for each, plus two more *control* days. Stimuli 2 (*darkness*) and 6 (*live monkey*) were not used in Experiment II.

C. RESULTS

1. *Experiment I*

All *S*s showed a trend of increasing door-pushing behavior throughout test sessions. The increase in responsiveness (cf. 5, 7) is illustrated by the mean total time pushing the door open during the first hour (before any stimulus change occurred). This value rose from a mean hourly rate of 318 seconds on test Sessions 1–5 to a rate of 678 seconds on Sessions 10–14. This difference is significant for the group at the .01 level. All *S*s were tested for 14 days and two were continued to 17 and 20 test days respectively. The greatest increase occurred between the 5th and 10th day in all *S*s.

The increasing amount of exploratory activity through test sessions was primarily the result of an increase in the average duration of each response. Only one *S* showed an increase in the frequency of responses per hour through test sessions (No. 31), while all increased total response time. The mean response duration for all *S*s was 7.1 seconds on the first five days and 11.7 seconds on the last five.

Observations made during test sessions by *E* via a small aperture near the test box light indicated that *S*s were not exploring visually during all the response time. They appeared to spend more than half the response time manipulating the door, chewing on the sill, and exploring tactually what they could not see through the open door. Very commonly they sat near the door holding it open with one hand and moving it against the spring resistance, and only occasionally turning to look through the opening. These informal observations suggested that this sort of activity increased through the period of testing and probably accounted for most of the rise in total response time.

Evaluation of the effects of the several external stimulus conditions on this behavioral measure is rendered difficult by the shifting baseline and the confusion of manipulatory with visual exploratory activity. Test hour response times were similar to first hour scores in showing great variability, large individual differences, and the above noted trend of increase. The limited number of presentations of each stimulus makes averaging raw response duration data impossible. If, however, each test day is treated separately as to the effect of stimulus change (first hour total response duration subtracted from test hour) some stimulus-correlated behavioral effects are seen. In Table 1 the data treated in this way for all *S*s are summarized. A cross in either the 'increase' or 'decrease' column means that the three test hour change scores for that stimulus condition showed no overlap with the five test hour change scores obtained on control days. This is thought to be a rather stringent criterion for the presence of an effect. No effects were disclosed from consideration of frequency of response scores that were not evident from this analysis of duration scores.

TABLE 1

EFFECT ON DOOR PUSHING BEHAVIOR OF SIX STIMULUS CONDITIONS: (Experiment I)

Stimulus	S	Increase	Decrease	No Change
Darkness	29			X
	31		X	
	32		X	
	74			X
Photograph	29		X	
	31	—	—	—
	32		X	
	74		X	
Live Monkey	29	—	—	—
	31	X		
	32	X		
	74	X		
Triangle	29			X
	31			X
	32			X
	74	—	—	—
Moving triangle	29	—	—	—
	31			X
	32			X
	74	—	—	—
Novelty	29	—	—	—
	31			X
	32			X
	74		X	

It became apparent that the only stimulus which produced a substantial and sustained increase in response time was No. 6, *live monkey* (cf. 5). All of the three *S*s exposed to this stimulus showed a marked increase in the time spent holding the door open during the test hour compared with the first or standard hour on every presentation. The actual proportion of the test hour spent holding the door open averaged 56 per cent in No. 32, 42 per cent in No. 31, 11 per cent in No. 74. External darkness, a situation presumed to be visually nonrewarding, produced a marked decrease in amount of responding in two *S*s and no substantial change in the other two.

The photograph of a fearful monkey produced a striking drop in amount of responding in all three *S*s exposed to it. This somewhat unexpected effect was accompanied by some evidence of excitement and conflict in the test monkeys, especially on the first presentation. Cage shaking, agitated approach-avoidance behavior toward the door, immediate shutting of the door when the screen came into view and unusual vocalizations were noted. In general, *S*s acted as they had during initial training to respond.

Even more unexpected was the change in response rate observed immediately at the start of the third hour (which began with *S*'s first response after the stimulus was removed) on those days when *photograph* or *darkness* was the test stimulus. A sharp increase, often reaching the highest level of the day, occurred, and usually lasted for 10 minutes or less. The results of plotting the average change from the rate established during the first hour by 10-minute periods are shown in Figure 2. The data from each test day were converted into this form as follows: the first hour total was divided by 6 to obtain an estimate of baseline rate; this was subtracted from the total response time in each 10-minute period following, and these scores averaged for the several stimulus conditions. It is clear from Figure 2 that a "rebound" effect was observed in every instance of a significant decrease in rate during the test hour. Our direct observations of *S*s during these periods convinced us that both the depression and "rebound" effects were consequences of increased fear, and that *photograph,* to a lesser extent *darkness,* and in one case *novelty,* were anxiety arousing stimuli.

Data obtained on control days are included in Figure 2 for comparison, and give an indication of changes in responsiveness *within* test sessions when no stimulus changes occurred. All *S*s responded more in the last hour than in the first, although the difference is not significant due to the high day-to-day variability noted above.

That the presence of movement in the visual field is no more rewarding than a stationary stimulus is illustrated by the results with Stimuli 3 and 4. The failure to observe a higher rate of responding when the triangle was moving about in its orbit argues against a simple explanation of visual exploratory behavior based on a reinforcing property of shifting visual patterns.

Locomotor activity as defined by frequency per hour of counts on the capacity relay revealed no significant correlation with frequency or duration scores on the door mechanisms. Two *S*s (Nos. 29 and 31) showed a moderate increase in first hour locomotor activity through test sessions paralleling

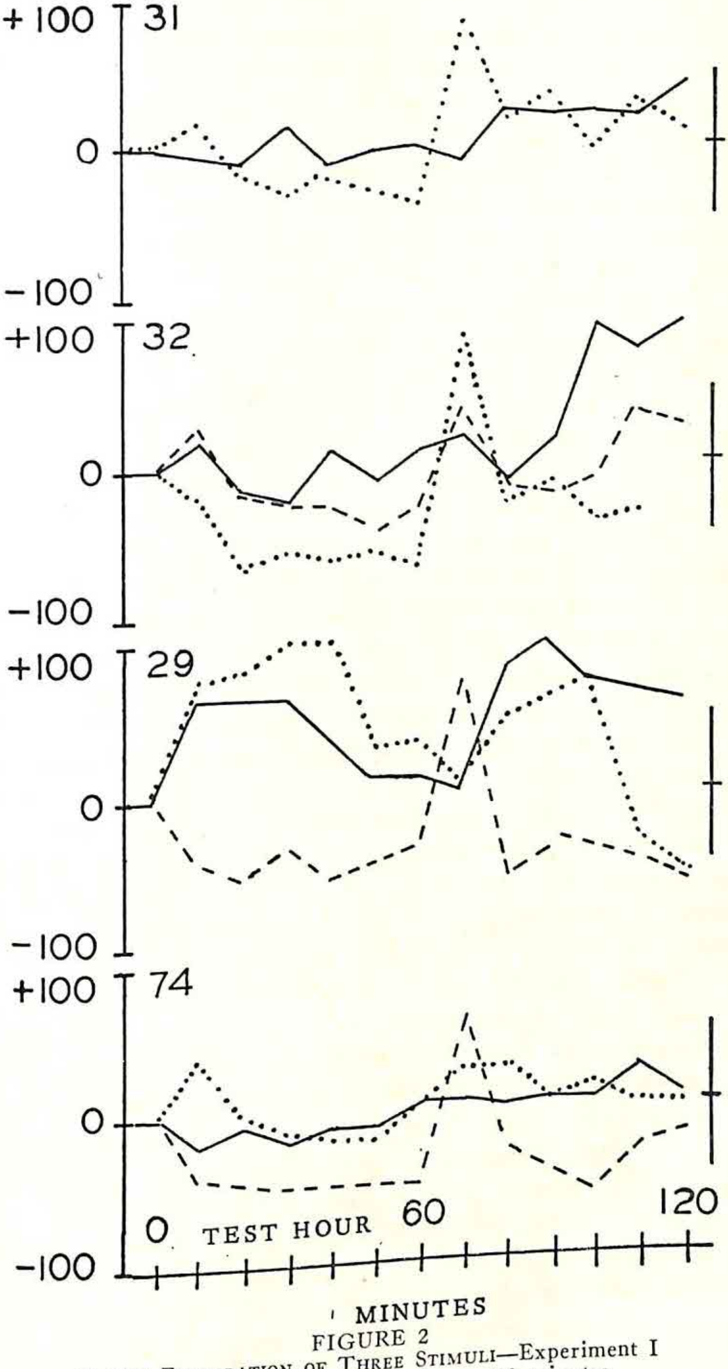

FIGURE 2
VISUAL EXPLORATION OF THREE STIMULI—Experiment I
0 = Rate of First Hour Exploration in seconds per 10 minutes
Solid line = Control; Dashed line = Photograph; Dotted line = Darkness

the rise in exploratory-manipulatory behavior on the door. Monkey 32 showed a moderate decrease through testing, and No. 74 maintained a very low rate. There is a suggestion that in all *S*s (except No. 74 whose rate remained low under all conditions) stimulus *live monkey* was associated with a drop in counts per hour and stimulus *photograph* was associated with an increase, but the differences are not significant. These data are consistent with observations of increased agitation of *S*s during exposure to *photograph,* and of preoccupation with the door during exposure to *live monkey* (sitting by the door and opening it whenever the object monkey vocalized).

2. *Experiment II*

Our direct observations of *S*s convinced us that response times recorded on the peephole mechanism reflected visual exploratory behavior almost exclusively. Some monkeys managed to sit for a time in a position which interrupted the light beam, a few placed their hands over the lenses, but only briefly and when they were unaccustomed to the test box. Responses on the door, on the other hand, especially in the naïve *S*s who had never experienced a visual reward via the door, must be primarily a measure of manipulatory behavior. There was no indication that the monkeys were trying to escape or to explore visually in the baffle. It was obvious that response rate on the door was only a partial measure of manipulatory activity. All portions of the test box were exhaustively manipulated and chewed by *S*s throughout test sessions.

Visual exploration of the standard stimulus during the first hour was relatively constant throughout test sessions, showing no significant increase or decrease (dashed line, Figure 3). The relatively low rate on the first three sessions is attributable to the scores of the two experienced *S*s, who were slower to learn to use the peephole than the naïve animals. The mean first hour response time for all six *S*s was 151.3 seconds ($SD = 110$). Individual differences in amount of exploring were not very great under the standard condition. First hour mean response times for the several *S*s ranged from just under one minute to about 4½ minutes. The average duration of exploratory responses was remarkably constant throughout testing for each *S*. Values for individual *S*s ranged from 2.4 to 5.5 seconds average response duration based on all first hour scores. A moderate increase in visual exploration *within* test sessions on control days was seen in four of the six *S*s in Experiment 2. For all *S*s the difference between the first and second 90-minute periods on control days was significant at the .03 level (illustrated by the solid line, Figure 4). This effect, which was observed in Experiment 1 also, seems to support the position that a "boredom drive" was operating in our *S*s in the test box (18). The alternative explanation that higher anxiety at the beginning of each test session inhibited exploration would be ruled out if our contention that mild anxiety increases exploration via the peephole is correct (see below). The further alternative that exploratory drive was increased by novelty (1, 16) seems unreasonable in this increasingly familiar test situation.

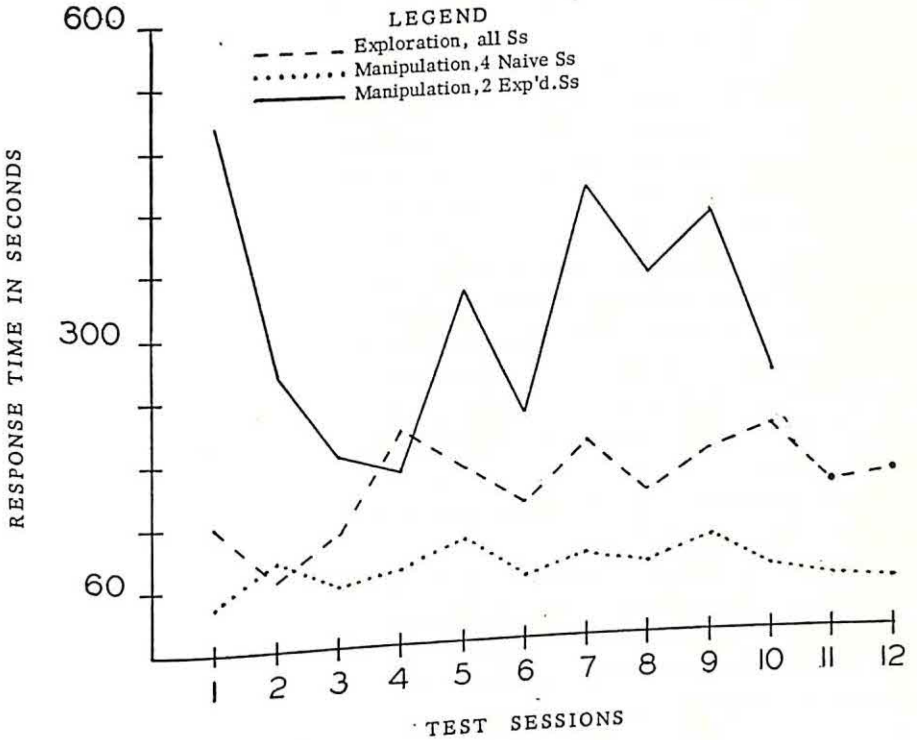

FIGURE 3
FIRST HOUR RESPONSE TOTALS—Experiment II

Response time on the door was also quite constant through test sessions, but showed wide individual differences which may be related to the previous experience of the two *S*s from Experiment I. First hour manipulation totals for four naïve and two experienced *S*s are given in Figure 3 for successive test sessions. We did not observe an increase in manipulation scores as might have been predicted from the results of Experiment I. The two experienced *S*s (Nos. 29 and 32) appeared to extinguish the door pushing response during their first four days and then to begin another trend of increasing activity (solid line, Figure 3). Whether the high rate of door pushing during Days 5 to 10 for these two *S*s is related to their earlier experience with the door or to individual differences in strength of the manipulatory motive is uncertain. We tend to favor the latter explanation. These two *S*s were the only ones on the average to give higher daily response times on the door than on the peephole. Among the four naïve *S*s mean total response duration ranged from about 40 seconds to about 100 seconds during the first hour. Although the naïve *S*s consistently displayed more interest in the peephole than in the door, the amount and persistence of manipulatory behavior is striking. None showed any regular decline after more than 30 hours of testing (cf. 13). We observed no significant correlations between amount of manipulation and amount of visual exploration during standard condition hours or the various test condition hours.

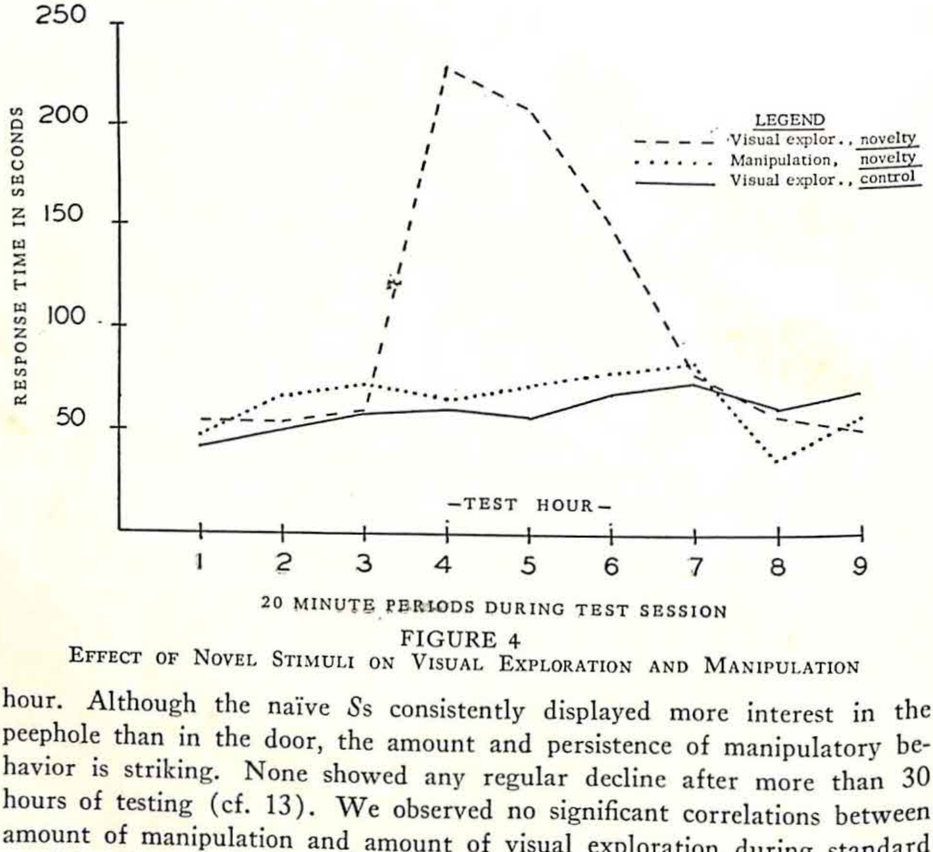

FIGURE 4
EFFECT OF NOVEL STIMULI ON VISUAL EXPLORATION AND MANIPULATION

The results of exposing *S*s to the various test stimuli clearly demonstrated that none of the individual slides had any significant effect on visual exploration either during or immediately after exposure. The regularly changing sequence of slides, however, which we have called *novelty*, produced a striking and sustained increase in visual exploration. Evidence of the reinforcing effect of *novelty* is adduced in Figure 4. The more reliable measure of baseline visual exploration obtained in Experiment II enabled us to plot mean duration scores directly. In all six *S*s the mean response rate (seconds per hour) during the three hours' exposure to novelty was significantly greater than that seen during the comparable control hours. Expressed in the relative units described above (test hour total minus first hour total each day) the effect on behavior of this sequence of projected images was greater by far than any effect seen in Experiment I. Test change scores for *novelty* averaged more than three times the corresponding *control* scores and showed no overlap in any *S*. For all six *S*s, the *control* hour mean response time was 189 sec. (*SD* 115) and the *novelty* hour mean response time was 587 sec. (*SD* 218). This difference is significant beyond the .01 level.

There was a suggestion that the reinforcing property of visual exploration of the slides was diminishing with increasing exposure. In five of six *S*s, the test change score for the last *novelty* day was the smallest of the three such scores, although still, as mentioned above, very much greater than any *control* score. The *S* from Experiment I which had been exposed to *novelty* for a total of three hours in that design was still responding at a rate nearly double the first hour rate at the end of the third test hour in Experiment II (sixth hour of exposure altogether). No attempt was made to investigate the recovery time of the reinforcing effect of *novelty*. The delay since the last previous exposure to *novelty* varied among *S*s from about 48 hours to more than a week, and no systematic effect on exploratory behavior was noted.

It is surprising in view of the results of Experiment I that exposure to the photograph of the fearful monkey had no significant effect on exploratory behavior. No depression of response rate during the test hour, and no "rebound" phenomenon or sharp increase in rate when the stimulus was removed, were seen. In fact, two naïve *S*s showed a moderate increase in exploration during the 60 minutes of exposure to the stimulus. These animals evidenced mild agitation when first exposed to the *photograph* stimulus.

Finally it may be noted that manipulatory activity as defined by opening a door leading to a dark chamber can persist relatively unaffected during wide shifts in the amount of visual exploration (dotted line, Figure 4). Whereas all *S*s displayed a striking increase in visual exploration to *novelty*, none showed any reliable change in response rate on the door during the same period. This again appears to argue that no simple relationship exists between the two kinds of behavior.

D. Discussion

These experiments were undertaken with the aim of illuminating the motivational basis of exploratory behavior in monkeys. It appears to us that several factors are contributing to the exploratory motive, for we observed both increased and diminished exploration of certain "novel" stimuli in our two experiments.

The only important difference between Experiment I and Experiment II of which we are aware lies in the nature of the response by which monkeys confined in a test chamber gained visual access to the external environment. Several observations convinced us that a great difference existed for the monkey between opening a door, thereby betraying his presence and exposing himself to possible harm, and looking through a small hole in the side of the test box. First, initial learning of the two responses was strikingly different. The *S*s in Experiment I showed a reluctance to put their hands through the door opening (as was necessary to hold the door open) and had to be deliberately trained to do so. Those monkeys which appeared to have a lower degree of excitability and fearfulness in their general behavior overcame this reluctance sooner than others, but it was still a hindrance to rapid learning. The *S*s in Experiment II, on the other hand, used the peephole immediately and required no special inducements to do so. Three of the four naïve *S*s explored more during their very first hour in the test box—a time when we would expect "boredom" to be low and anxiety to be high—than during any of the following 11 standard condition hours. Second, behavior toward *E* was different. In Experiment I all *S*s refused to open the door when they had seen *E* outside the test box (making repairs, usually) or knew he was there. The *S*s in Experiment II, however, who seemed otherwise to be about as fearful, watched *E* closely through the peephole. This difference between groups was also seen in the individual *S*s which were tested in both designs. Third, exploration of what appeared to be a mild anxiety provoking stimulus (*photograph*) was very different in the two experiments, as was noted in the results section above. The importance of fear in the determination of these responses convinced us that anxiety reduction must be included in any account of the source of exploratory behavior. Three hypotheses about the rôle of anxiety in our two experimental situations are therefore proposed. In these, anxiety reduction is viewed as a reinforcing event.

1. Visual exploration of a potentially dangerous environment is an anxiety reducing response.

2. Concealment in a potentially dangerous environment is an anxiety reducing response, and exposure is an anxiety increasing response.

3. When anxiety level is high, novel stimuli are more apt to arouse anxiety than they are when anxiety is low.

These propositions assist in comprehending some of our observations, as follows: the slower initial learning and evidence of conflict behavior in Experiment I are explained by Propositions 1 and 2, which suggest that exposure and visual exploration have opposite effects on anxiety. In our Experiment II exposure was minimal and the effect on anxiety of exploration via the peephole was wholly or predominantly in the direction of reduction. We failed to observe, in fact, any situation involving change in the external environment which *S*s in Experiment II preferred not to explore. A number of stimuli, however, were avoided by *S*s in Experiment I. We find it reasonable to regard the behavior of *S*s here as a resolution between the anxiety reducing response of concealment from "danger" and the anxiety reducing response of exploring "danger." The result was a drop in response rate during the period of stimulus presentation (reducing "exposure" anxiety), and an increase in response rate after stimulus removal (reducing "presence of danger" anxiety). Proposition 3 might explain the neutral or inhibiting effect of stimulus *novelty* in Experiment I, when results of Experiment II demonstrated that *novelty* greatly increased exploration. That stimulus change can arouse fear under some circumstances is a well documented but entirely descriptive conclusion (1).

Several stimuli appeared to us to fit the description "anxiety arousing" for our *S*s. We do not know what factors were common to all, or indeed whether one or several kinds of fear were operating. In the case of the photograph of the fearful monkey, it is tempting to speculate that the expression and posture (restrained) of the depicted animal were important factors, but controlled evidence is lacking. External darkness was less potent in its effect on exploratory behavior, but appeared similar to *photograph* in direction of effect. It is possible that fear of darkness in monkeys is related to their vulnerability to predators in the wild state at such times.

While evidence for the important rôle of anxiety reduction is compelling, we do not feel that it is alone a sufficient explanation of the exploratory motive. The tremendous increase in amount of exploration, the sustained nature of this interest, and the absence of other signs of correspondingly heightened anxiety lead us to conclude that "novelty," a term generally referring to dissimilarity in a sequence of stimuli, can have a powerful reward effect in the monkey. When sensory and behavioral homogeneity have been maximized, responses leading to a change are reinforced. When anxiety level is low, novelty (at least some kinds of dissimilarity) reinforces without change in the level of anxiety. We prefer to think of the reward effect of novelty as the product of an externally aroused drive which is related to environmental variables such as sensory homogeneity, confinement, isolation, etc. Behavior sequences occurring when novelty is not present are strongly reinforced if they result in novelty (cf. 10). Homogeneity is avoided by means of seeking novel stimulation and the terms "homogeneity drive" or "boredom drive" may be useful and consistent with other terminology (18). Anxiety is another source of exploratory motivation and may obscure the relationship between homogeneity and exploratory behavior. We do not feel that conceptualizing boredom as an anxiety-like state resolves this difficulty.

Manipulatory behavior appears to be yet another factor complicating the picture of visual exploration. It is remarkably persistent without stimulus change and seems to continue at a relatively stable rate during wide shifts in the rate of visual exploration (which imply fluctuations in the strength of the "boredom drive"). Individual differences among monkeys were observed to be large, apparently greater than those seen in the case of visual exploration. The confusion of visual exploration and manipulatory activity is regarded by us as a disadvantage of the method used in Experiment I, separate from and in addition to the greater rôle of anxiety discussed above. It is interesting to note that reinforcement from tactual exploration (manipulation) depends at least in part on certain attributes of the manipulanda themselves (11, 17) whereas visual exploration requires the property of stimulus change for its reward effect (when anxiety is low). Determination of the kinds and amounts of stimulus change which are maximally reinforcing is an important question for further study.

E. SUMMARY

Two experiments are reported dealing with visual exploration of a number of controlled stimuli by monkeys. The two experiments were alike except for the nature of the instrumental response made by the monkey to obtain visual access to a richer environment. In the first, a swinging door four inches square had to be pushed open, exposing the monkey while permitting him to view the environment outside his test box. In the second, a small peephole was provided through which the monkey could look at will while maintaining nearly complete concealment.

Differences in early learning, in behavior while *E* was nearby, and in amount of exploration to several stimuli which appeared to evoke mild fear in the monkeys suggest that anxiety reduction is an important factor in the

exploratory motive. Three hypotheses about the rôle of anxiety reduction in exploratory behavior are proposed. Predictions from these hypotheses appear to parallel some of our experimental results.

Anxiety reduction, however, is not a sufficient explanation of the exploratory motive. A series of photographic slides projected into the field visible through the peephole and automatically shifted every 90 seconds produced a striking increase in amount of exploration which was sustained over long periods of time. Accordingly it is concluded that under conditions of relative sensory deprivation novel stimuli have the property of reinforcing responses made when novelty was not present. This reward effect is thought to be independent of anxiety reduction. We find our results best explained by a theory such as that proposed by Myers and Miller (18), which regards isolation and stimulus homogeneity as drive producing events, and exploration of novelty as a drive reducing response.

REFERENCES

1. BERLYNE, D. E. Novelty and curiosity as determinants of exploratory behavior. *Brit. J. Psychol.*, 1950, **41**, 68-80.
2. BERLYNE, D. E., & SLATER, J. Perceptual curiosity, exploratory behavior, and maze learning. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 228-232.
3. BEXTON, W. H., HERON, W., & SCOTT, T. H. Effects of decreased variation in the sensory environment. *Canad. J. Psychol.*, 1954, **8**, 70-76.
4. BUTLER, R. A. Discrimination learning by rhesus monkeys to visual-exploration motivation. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 95-98.
5. ———. Incentive conditions which influence visual exploration. *J. Exper. Psychol.*, 1954, **48**, 19-23.
6. ———. The effect of deprivation of visual incentives on visual exploration motivation in monkeys. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 177-179.
7. ———. Discrimination learning by rhesus monkeys to auditory incentives. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 239-241.
8. BUTLER, R. A., & ALEXANDER, H. M. Daily patterns of visual exploratory behavior in the monkey. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 247-249.
9. BUTLER, R. A., & HARLOW, H. Persistence of visual exploration in monkeys. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 258-263.
10. CHAPMAN, R. M., & LEVY, N. Hunger drive and reinforcing effect of novel stimuli. *J. Comp. & Physiol. Psychol.*, 1957, **50**, 233-238.
11. CHO, J. B., & DAVIS, R. T. Preferences of monkeys for objects other than food. *Amer. J. Psychol.*, 1957, **70**, 87-91.
12. DAVIS, R. T., SETTLAGE, P. H., & HARLOW, H. F. Performance of normal and brain operated monkeys on mechanical puzzles with and without food incentives. *J. Genet. Psychol.*, 1950, **77**, 305-311.
13. HARLOW, H. F., HARLOW, M. K., & MEYER, D. R. Learning motivated by a manipulation drive. *J. Exper. Psychol.* 1950, **40**, 228-234.
14. HARLOW, H. F., & McCLEARN, G. E. Object discrimination learned by monkeys on the basis of manipulation motives. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 73-76.

15. Montgomery, K. C. The rôle of the exploratory drive in learning. *J. Comp. & Physiol. Psychol.*, 1954, **47**, 60-64.

16. ———. The relation between fear induced by novel stimulation and exploratory behavior. *J. Comp. & Physiol. Psychol.*, 1955, **48**, 254-260.

17. Murray, E. J. The effects of hunger and type of manipulandum on spontaneous instrumental responding. *J. Comp. & Physiol. Psychol.*, 1953, **46**, 182-183.

18. Myers, A. K., & Miller, N. E. Failure to find a learned drive based on hunger: evidence for learning motivated by "exploration." *J. Comp. & Physiol. Psychol.*, 1954, **47**, 428-436.

19. Whiting, J. W. M., & Mowrer, O. H. Habit progression and regression—a laboratory study of some factors relevant to human socialization. *J. Comp. & Physiol. Psychol.*, 1943, **36**, 229-253.

Department of Physiology
Yale University School of Medicine
333 Cedar Street
New Haven 11, Connecticut

The Journal of Genetic Psychology, 1959, **94**, 199-203.

BEHAVIOR OF THE OPOSSUM IN THE FINK ARROW MAZE*

Department of Psychology, University of Georgia

W. T. JAMES[1]

A. PROBLEM

In a program of research on behavior of the opossum at the University of Georgia the last few years (2, 3, 4), the animal has shown surprising adaptability. Apparatus devised by various experimenters have been used in the tests, including the Guthrie-Horton Puzzle box, two-door discrimination box, lever operation to open doors, the Skinner apparatus and conditioning procedures.

Recently Fink (1) published an extensive study on the behavior of 9 species of animals in a simple maze learning situation. The interest in the present experiment was comparing the behavior of the opossum with those studied by Fink, using exactly the same apparatus and procedure.

B. APPARATUS

The apparatus consisted of a choice compartment with 4 parallel alleys leading off one side. The choice area was 18 inches by 4 feet by 2 feet high. Each alley was 8 inches wide. The two outside alleys were 28 inches long, while the two inside alleys were 14 inches long. At the end of each alley there was a 90 degree turn toward a center partition running through the center of the structure. This gave a partition behind which food could be placed so it could not be observed from the front of the alley. The whole maze was covered by wire and a hinged top so the animals could be removed from the alleys after each test. A two foot square start box was attached to the choice area with a sliding door for releasing the animal into this area before each test.

C. PROCEDURE

The problem set for the animal was the location of food. The animal was placed in the start box just before the test each day at regular feeding time. Only one run was made each day. As a rule, food was placed in Alley 3 for the first tests. The animal explored until it found food in Alley

---

* Received in the Editorial Office on February 11, 1958.

[1] This research was supported by the Graduate School of the University of Georgia.

3. After eating the meat the animal was replaced in the home cage until the test the next day. This process was repeated until the animal entered the alley for 10 straight times without a mistake. The food was then shifted to Alleys 1, 2, and 4 and repeated. The time required to reach food each day was recorded along with a tracing of the path taken by the animal in finding food. Since the opossum is highly sensitive to odors, attempts were made to mask the olfactory cue for the correct alley. To guard against this, food (meat) was placed on the wire above each incorrect alley. By this procedure odor permeated the whole maze and the animal was forced to rely on spatial location to find food.

D. Animals

Five adult animals were used in the experiments. They were born and reared in the psychological laboratory. Since they were fed at regular intervals once a day, and were accustomed to eating in the presence of the *E*, they would go to work immediately in the maze tests. None of the animals exhibited the 'passive defense' behavior so often found in wild opossums. They were not disturbed by strangers or any form of extraneous noises in the laboratory.

E. Results

The plan used by Fink in scoring the tests has been used in this study. The first score is the *P* value, the Performance Quotient. This is the sum of trials required for learning all four alleys. The next score was the *T* value, or Trial Quotient. This is the same as the *P* value except that it excludes the correct trials among those required for learning the maze. Thus if an animal entered the correct alley early in the test by chance, it was not included among the *T* values. This value does not include the last 10 runs which fit the criterion of learning. The third score was the *AQ*, or Alley Quotient. Many of the animals studied by Fink would enter part way into the correct alley and withdraw. In some cases they would partially enter a wrong alley and withdraw. The *A* value includes all these partial choices which are called entrance errors. The opossum did not exhibit this form of behavior. In learning new alleys, however, in many cases the animal would start out in the wrong direction, approach the wrong alley part way, then turn away just before entering. The *AQ* value includes all these orientation errors, which did not result in alley entrance. This form of behavior was especially shown in Animal 3 while learning the last alley in the series.

On the basis of the *T*, *P*, and *A* values, Fink calculated what is called the *T*-Quotient, the *P*-Quotient and the *A*-Quotient. All of these scores were determined in relation to those made by man in learning exactly the same form of finger maze while blindfolded. The *P*-Value for man, for example, was 14.8, and the *PQ* for man was 100. Other animals were related to the value of man. For example, the pig had a *P* value of 31. The *PQ* of the pig calculated in the following manner, $\frac{14.8}{31} \frac{\text{Man}}{\text{Pig}} = \frac{\text{PQ}}{100}$.

Thus $PQ = 1480/31 = 47$. This gives the learning of the pig in relation to that of man. The *T*-Value was calculated in the same manner. The *T*-Value for man was 9.2 trials, and the *TQ* was arbitrarily valued at 100. Thus *TQ* equals 920/T. The *A*-Value for man was 25.1. In calculating the *A*-Value of the animals the *AQ* equals 2510/A. The value of the opossum, and the other animals studied by Fink are shown in Table 1.

TABLE 1
PTA VALUES WITH *PQ*, *TQ*, AND *AQ* SCORES OF ANIMALS TESTED IN THE ARROW MAZE

Av. rank	Animal	*P*	*T*	*A*	*PQ*	*TQ*	*AQ*	Aq av.
1	Man	14.8	9.2	25.1	100	100	100	100
2	Opossum (2)	24	22	26	61	51.5	96	65
	Opossum (1)	27	20	29	54	46	84	61
	Opossum (5)	27	19	35	54	48	71.7	57.9
	Opossum (3)	29	28	39	47	32	64	64
	Opossum (4)	38	28	46	38.9	48.4	53	46.7
2	Pig	31.8	21.8	46.7	46.5	42.2	53.7	47.5
3	Dog	38.4	26.1	40.4	38.5	35.2	62.1	45.3
4	Goat	54	37	56	27.4	24.9	44.8	32.2
4.5	Chick	46.4	34.2	68.1	31.9	26.9	36.9	31.9
4.5	Rat	53.7	32.8	62.4	27.6	28	40.2	31.9
5	Cat	61.1	43.1	89.5	24.2	21.3	28	24.5
6	Rabbit	67	37	67	22.1	24.9	37.5	28.2
	Turtles below this level (see Fink).							

As observed from the table, the opossum falls between man and the performance of the pig in this type of behavior.

F. DELAY OF RESPONSE

The same general picture of delay of the response was exhibited by each animal. Rather than give a curve showing the average delay of all animals, a graph is made of the delay of Animal 3 in Figure 1. The animal begins with a delay of 232 seconds and ends on the last alley with a delay of two seconds. In general there is a rise in time as the goal is shifted from one alley to another. This delay decreases, however, toward the end of the tests.

G. Discussion of Results

There has been surprisingly little work done on the behavior of the opossum. This is the case in spite of the interest shown in the embryology and neurology of the animal and the folk lore regarding them. The general opinion among hunters and furriers is that the animal is particularly unintelligent. This is probably based on the animals inactive nature, and exhibition of 'passive defense' when disturbed. In view of this the results under experimental conditions have been surprising.

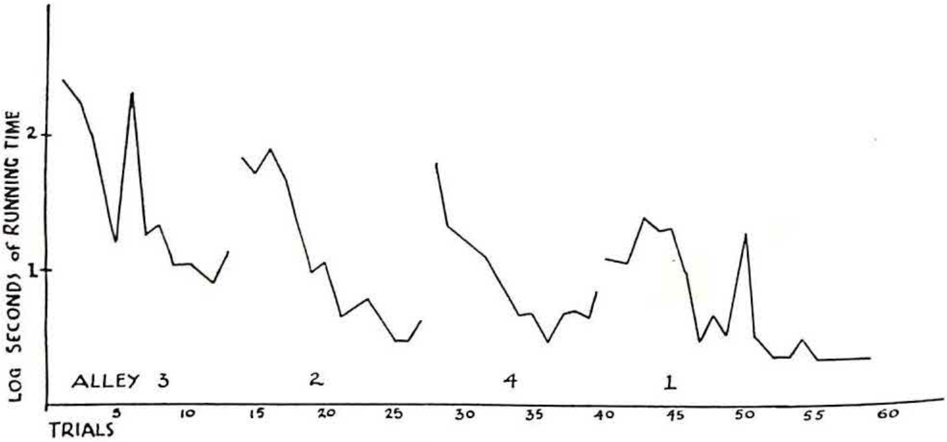

FIGURE 1
This Figure Shows the Log Seconds of Running Time for Each Trial as the Animal Learned All Four Alleys

Since the opossum belongs to the second order of mammals with a relatively simple cortex (5), it would seem that it should fall much lower in the scale in the Fink maze test. Apparently in regard to spatial location it performs as well as other animals. It is highly possible that this test deals only with a form of instinctive behavior and not with behavior which measures 'intelligence' in the usual sense. It may be that the greatest change in animals occurs when the level of symbolic representation is reached, and there is relatively little difference among animals below this, except in the range of sensitivity, and manipulative ability. There can be no doubt that man has underrated the opossum in any case.

H. Summary

In this experiment the behavior of five opossums has been compared with the nine species of animals studied by Fink in the Arrow Maze. Exactly the same method and procedures were used as in Fink's original study. In this test the animal learns the location of food alley ways going off a center vestibule. After the location in one alley is learned and run for 10 correct trials, food is then shifted to another alley. This process is repeated until all four alleys are learned. In this test the opossum falls below man and near the pig on the various scores. The performance was much better than expected on the basis of the poor cortical development of this animal.

REFERENCES

1. FINK, H. K. Mind and Performance. New York: Vantage Press, 1954.
2. JAMES, W. T. The behavior of the opossum in the Guthrie-Horton puzzle box. *J. Genet. Psychol.*, 1955, **87**, 203-206.
3. ———. An experimental study of the defense mechanism in the opossum with emphasis on natural behavior and its relation to mode of life. *J. Genet. Psychol.*, 1937, **51**, 95-100.
4. ———. Conditioned responses in the opossum. *J. Genet. Psychol.*, 1958, **93**, 179-183.
5. PAPEZ, J. W. Comparative Neurology. New York: Crowell, 1929.

Department of Psychology
University of Georgia
Athens, Georgia

The Journal of Genetic Psychology, 1959, **94**, 205-209.

THE EFFECTS OF LONG-TERM HOSPITALIZATION OR INSTITUTIONALIZATION UPON THE LANGUAGE DEVELOPMENT OF CHILDREN*

St. John's Episcopal Hospital, Brooklyn

Arthur D. Haggerty

A. Introduction

Modern hospital administration and clinical pediatric practice have completely adopted the premise that the hospital environment should be as much as possible a "home away from home" for the patient. Efforts to effect such a condition rather than impeding administrative or medical efficiency, as might have been initially protested, actually have enhanced the modern hospital's ability to rehabilitate both the physical condition as well as the emotional tone of a patient within a briefer period than before. Today very few hospitals fail to attempt combining clinical and administrative efficiency with hotel-like pleasantness since all believe that the emotional support of a patient is just as important as his surgical or medical care.

When we consider chronic care, particularly that of children, establishing an optimal environment becomes more difficult. Articles like that of Faust (6) indicate that this problem is recognized theoretically; one wonders, however, whether responsible officials are genuinely sensitive to the seriousness of the matter and thus are willing to put theory into practice. The purpose of this article is a demonstration that this problem must be solved unless children under chronic care are to return to their family physically improved but emotionally weakened.

Existing hospital, medical, and psychological literature is replete with indications that prolonged separation of a child from the family situation can be disastrous for the child's overall personality status. There is evidence that separation can affect something as basic as physical growth and rate of development as found by Bakwin (2) and Fried and Mayer (8). A parental type of fondling and handling is even known to affect such factors in experimental laboratory animals as reported by Weininger (16). Removing the child from parental association has also affected level of adjustment. During World War II in England Alcock (1) found that the trauma of bombing attacks did not disturb children nearly so severely as did evacuation

---

* Received in the Editorial Office on February 11, 1958.

with its separation from family ties. Similarly, Spitz (15) found a high correlation between psychopathic states in childhood and the absence of the emotional stimulation of the family and the influence of the mother. Fischer (7) studied a group of hospitalized six-month-old infants and concluded that there is a definite hospitalism syndrome which occurred in about one third of these children. She found that this condition was remedied only after the infants had been placed in a favorable family or family-like atmosphere. While most investigators feel that such an inhibiting force can ultimately be overcome, Goldfarb (9) expressed his belief that children who have spent much time in an institutional setting, particularly those placed there at an early age, could not improve even when later placed in a better environment since ". . . the extreme psychological deprivation in infancy produces a lag in mental growth which is maintained even under new conditions of enriched stimulation." Thus there can be no doubt about the inevitable possibility of children being adversely affected by prolonged hospitalization or institutional placement unless conditions that substitute for those found in the home are provided.

B. The Study

The purpose of this study is an evaluation of the effects of early and prolonged institutional types of placement upon a child's language development and thus eventually, through the disruption of this important function, the evolution of personality disorder. Data were gathered over a five-year period on 100 children who as clients of a specialized aftercare social agency had earlier in life spent a considerable amount of time in various hospital and child caring institution settings. The group's mean age was 12.7 years, and all were in the seventh grade of grammar school. The average age of separation from parents was 3.7 years of age with many having spent the first few years of their life in such situations. The average length of placement or hospitalization was 3.5 years.

In two previous studies the writer (11, 12) had investigated a similar group and found that although their general intellectual potential was average, they appeared retarded in functioning, particularly in language usage. That limitation in verbal expression is more than just a social handicap is advanced by the proponents of the Communication Theory in Psychiatry. Meerloo (13) feels that the "communicative act" is extremely important since "talking to each other is loving each other" and communication is the means whereby man fulfills his instinctive desire to reach other men's minds. It has also been further established by Moore (14) that children in placement have impaired communicability since they do not have so many experiences or materials to talk about as does the child in his own home. Neither do they have the continuously warm, stimulating communication experience with adults which is so necessary. Goldfarb (10) found that institutionalized children have significant degrees of personality difficulties and even in some respects resemble those afflicted with the severe disorder of Schizophrenia. It is also a well accepted fact that both early emotional deprivation and communication impairment exist in Schizophrenics. It is the hypothesis of this study that early and prolonged hospitalization or institutionalization can affect a child's personality development through the disruption of verbal communication ability.

The language of the subjects used was studied with the parts-of-speech analysis used by Boder (3) and by Ellsworth (5). To obtain language samples, oral responses were recorded from a sentence completion test and from stories given to the Thematic Apperception Test. The oral vocabulary subtest of the Wechsler Intelligence Scale for Children was also included, while the Rorschach Test was administered for personality evaluation. All verbal responses were transcribed verbatim with a recording instrument so that the samples obtained would be truly spoken language rather than written language where such factors as reading and writing ability could interfere.

C. RESULTS

The summary of the parts-of-speech test is portrayed in Table 1. Here it was found that the group studied had a formal language content that did not coincide with the pattern for their own grade level, but rather agreed more closely with a lower grade and even with the pattern for schizophrenics who are considered to have a "regressed" language facility. The norms for the fifth and seventh grades as well as those for the schizophrenic group were taken as controls from the study of Ellsworth (5).

TABLE 1

PERCENTAGES OF THE PARTS OF SPEECH FOR THE HOSPITAL-INSTITUTIONAL, SCHIZOPHRENIC AND CONTROL GROUPS

Group	Percentages of the parts of speech			
	Nouns	Pronouns	Verbs	Adjectives
5th Grade	27.6	19.2	30.5	21.9
7th Grade	26.8	18.1	30.0	25.1
Schizophrenics	27.5	20.1	30.9	22.0
Hospital-institution	27.2	19.0	30.3	23.1

Qualitative interpretation of the vocabulary responses was recorded for the subjects as done in the study of Chodorkoff and Mussen (4), comparing normals to schizophrenics. The institutionalization group proved surprisingly similar to schizophrenics in their selection of functional and example types of definitions as against the higher conceptual abstractions of normals.

Both quantitative and qualitative evaluation of the Rorschach results suggested that the group studied deviated from the hypothetical optimum in a reduced grasp of reality, less productive imagination, more anxiety, and affective lability as well as lessened ability to form deep relationships with others.

D. Discussion

It appears from the results of this study that early and prolonged hospital or other institutional experience can damage personality integration and can lead to an inhibition of proper communicability. There exists the great possibility that impaired communicability can lead into distorted personality growth by forcing the individual to withdraw more within himself since he is unable to communicate on an adequate social-emotional level with those about him. Thus his vital inner psychological and social needs become thwarted, progressing into the pathological development which can be seen in the victim of separation from the warmth and security which only an adult totally interested in the child can give.

Throughout much of the material obtained there has been apparent a thread of similarity between the child who is a product of lengthy institutionalization and the schizophrenic. Both are considered victims of early emotional deprivation. The language and conceptual impairment in schizophrenia has usually been considered the result of a psychotic condition which has interfered with thought processes. The evidence presented here, however, raises the question whether, just as with the communication-deficient separated child who has no psychotic thought disturbance, the schizophrenic during his early emotional and social deprivation experience may have had a similar but even more limited communication retardation which ultimately led into a withdrawal from close psycho-social contacts since he found that he could communicate only poorly or not at all and thus had many of his basic needs unmet. Such a withdrawing from and even hostility toward the apparently cold and unaccepting environment could very easily lead into the self-satisfying autistic thought life of the schizophrenic. Thus verbal communicative disability may not be so much a symptom of this type of psychosis as it is a condition which anticipates the disorder and actually helps cause its evolution. Impaired verbal ability may lead into autistic thought processes rather than the reverse as is most commonly assumed. Preventive measures, particularly in institutional settings, can be prepared to avoid such a reaction in the child under prolonged care; remedial approaches in addition to the usual psychotherapeutic techniques may more profitably stress improved verbal facility not only to project emotional content, but also to reestablish social-verbal confidence and skill.

E. SUMMARY

Over a period of five years data were gathered on 100 seventh grade children who had spent a prolonged period in some type of hospitalization or other institutionalization. Their formal speech organization as well as overall personality characteristics resemble closely that of schizophrenics. It is suggested that language impairment may contribute to the development of schizophrenia rather than just being one of its symptoms.

REFERENCES

1. ALCOCK, T. Conclusions from psychiatric work with evacuated children. *Brit. J. Med. Psychol.*, 1948, **21**, 181-184.
2. BAKWIN, H. Emotional deprivation of infants. *J. Pediat.*, 1949, **35**, 512-521.
3. BODER, D. Adjective-verb quotient; a contribution to the psychology of language. *Psychol. Rec.*, 1940, **3**, 310-343.
4. CHODORKOFF, B., & MUSSEN, P. Qualitative aspects of the vocabulary responses of normals and schizophrenics. *J. Consult. Psychol.*, 1952, **16**, 43-47.
5. ELLSWORTH, R. The regression of schizophrenic language. *J. Consult. Psychol.*, 1952, **15**, 387-391.
6. FAUST, O. Stop scaring the children; hospitalization need not be a traumatic experience. *Mod. Hosp.*, 1953, **80**, 94-96.
7. FISCHER, L. Hospitalism in six-month-old infants. *Amer. J. Orthopsychiat.*, 1952, **23**, 522-533.
8. FRIED, R., & MAYER, M. Socio-economic factors accounting for growth failure in children living in an institution. *J. Pediat.*, 1948, **33**, 444-451.
9. GOLDFARB, W. Effects of psychological deprivation in infancy and subsequent stimulation. *Amer. J. Psychiat.*, 1945, **102**, 18-33.
10. ———. Rorschach test differences between family-reared, institution-reared, and schizophrenic children. *Amer. J. Orthopsychiat.*, 1949, **19**, 624-633.
11. HAGGERTY, A. D. The intellectual functioning of a post-institutional group. *J. Genet. Psychol.*, 1953, **83**, 303-306.
12. ———. The use of the altitude with a post-institutional group. *J. Genet. Psychol.*, 1954, **85**, 341-344.
13. MEERLOO, J. Conversation and Communication: A Psychological Inquiry into Language and Human Relations. New York: Internat. Univ. Press, 1952.
14. MOORE, J. Speech content of selected groups of orphanage and non-orphanage preschool children. *J. Exper. Educ.*, 1947, **16**, 122-133.
15. SPITZ, R. The effects of institution residence on inquiry into the genesis of psychopathic states in early childhood. *Rev. Franc. Psychoanal.*, 1949, **13**, 397-425.
16. WEININGER, O. The effects of early experience on behavior and growth characteristics. *J. Comp. & Physiol. Psychol.*, 1956, **49**, 1-9.

St. John's Episcopal Hospital
480 Herkimer Street
Brooklyn 13, New York

The Journal of Genetic Psychology, 1959, **94**, 211-215.

THE RELATION BETWEEN PRE-IRRADIATION VOLITIONAL ACTIVITY AND POST-IRRADIATION DEATH ORDER IN THE RAT*

Department of Psychology, University of Texas

Arnold A. McDowell and W. Lynn Brown[1]

A. Introduction

Kimeldorf and Jones (3) have demonstrated that forced exercise alters the mortality rate of the rat when given following irradiation, but that forced exercise prior to irradiation has no significant effect upon mortality. Jones *et al.* (2)have reported deficits in volitional activity of the rat following x-irradiation.

The present experiment by way of contrast was designed to study the relationship between pre-irradiation volitional activity and post-irradiation death order.

B. Materials and Methods

1. *Group I*

In the initial phase of the experiment 14 young adult male Sprague-Dawley rats were caged separately and maintained on a 24-hour food-deprivation schedule for 10 days. Each rat had free access to water and was given five pellets of Purina Laboratory chow daily at 4:30 P.M.

After the establishment of the food-deprivation schedule, the volitional activity of each rat during each of the eight hours from 8:00 A.M. to 4:00 P.M. was determined over a one-day period by use of the standard activity wheel. Each rat had free access to the wheel from the living cage of the apparatus.

The rats were then irradiated with 1200 *r* cobalt^{60} at 1200 *r* per minute in the high-level Co^{60} facility (1) at this laboratory. The rats were placed in individual containers and rotated during irradiation to produce radiation symmetry.

Following irradiation, the "death order" was recorded over a 30-day

---

* Received in the Editorial Office on February 12, 1958.

[1] This work was conducted at the Radiobiological Laboratory of The University of Texas and the United States Air Force, Austin, Texas, and was supported (in part) with funds provided under Contract AF 41(657)-149 and the School of Aviation Medicine.

period and correlations were established between pre-irradiation volitional activity measures and post-irradiation death order.

2. *Group II*

Next, using the same procedure as outlined above, pre-irradiation volitional activity measures were obtained for each of 10 young male black non-agouti hooded rats. On the basis of the correlations established with the first group of rats, death order predictions and a prediction of lesser susceptibility in the hooded rat were made.

The rats were then exposed to the same dosage of irradiation as was used with the first group of rats. Death order was noted. Since seven of the 10 rats survived this initial dosage, they were subjected to the same dosage 60 days after their first exposure. This second dosage was necessary to establish death order for the group.

C. RESULTS

Using the rank order correlation method (5) on the data for Group I, the Sprague-Dawley rats, it was found that mean rate of volitional activity correlated significantly with death order beyond the 1 per cent confidence level ($r' = .662$). It was also found that median rate of volitional activity correlated significantly with death order beyond the 5 per cent confidence level ($r' = .572$). Those rats with the higher mean or median rates died sooner than those with the lower rates, rate being defined as the mean or median revolutions of the activity wheel per hour during those hours in which the rats recorded revolutions of the wheel.

Figure 1 compares the two groups with respect to mean revolutions per hour for the eight successive hours prior to feeding. Using the Mann-Whitney U test (5) to compare the activity of the rats in the two groups, it was found that the activity of the rats in Group II, the hooded rats, was significantly less than the activity of the rats of Group I, the Sprague-Dawley rats, beyond the 1 per cent confidence level. The mean rates for the rats of Group II were determined to be significantly less than those for the rats of Group I beyond the 2½ per cent confidence level, the median rates to be significantly less beyond the 5 per cent confidence level. Since this was the case and since both mean and median rates had correlated significantly with death order in Group I, it was predicted that the rats of Group II would die less rapidly than the rats of Group I.

Figure 2 plots the cumulative percentage of deaths for the two groups during the first 30 days following irradiation. Statistical comparison of the two groups with respect to the number of deaths and number of survivals during this period yields a Chi-square value of 5.5312, which is significant beyond the 2 per cent confidence level. The hooded rats were less susceptible than the Sprague-Dawley rats.

Death order predictions were made for the rats of Group II on the basis of both mean and median rates. For the rats of Group II, mean rate was found to correlate significantly with death order beyond the 5 per cent confidence level ($r' = .612$). Median rate was found to correlate significantly with death order beyond the 1 per cent confidence level ($r' = .697$).

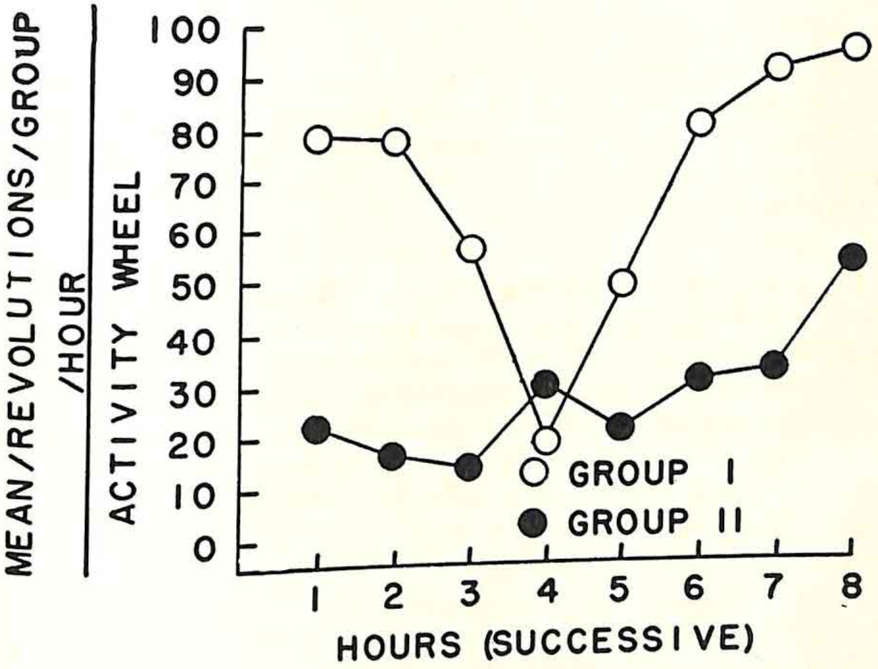

FIGURE 1
THE MEAN ACTIVITY WHEEL REVOLUTIONS PER GROUP PER HOUR FOR THE EIGHT SUCCESSIVE HOURS PRIOR TO FEEDING TIME

D. DISCUSSION

It is apparent from the results of the present study that individual differences in susceptibility of adult male rats to irradiation are related to differences in volitional activity rates, the rat with a higher rate being more susceptible than the rat with a lower rate. The present study also demonstrates the validity of pre-irradiation predictions of post-irradiation death order.

It seems probable, although post-irradiation deficits occur in volitional activity (2), that post-irradiation volitional activity correlates directly with

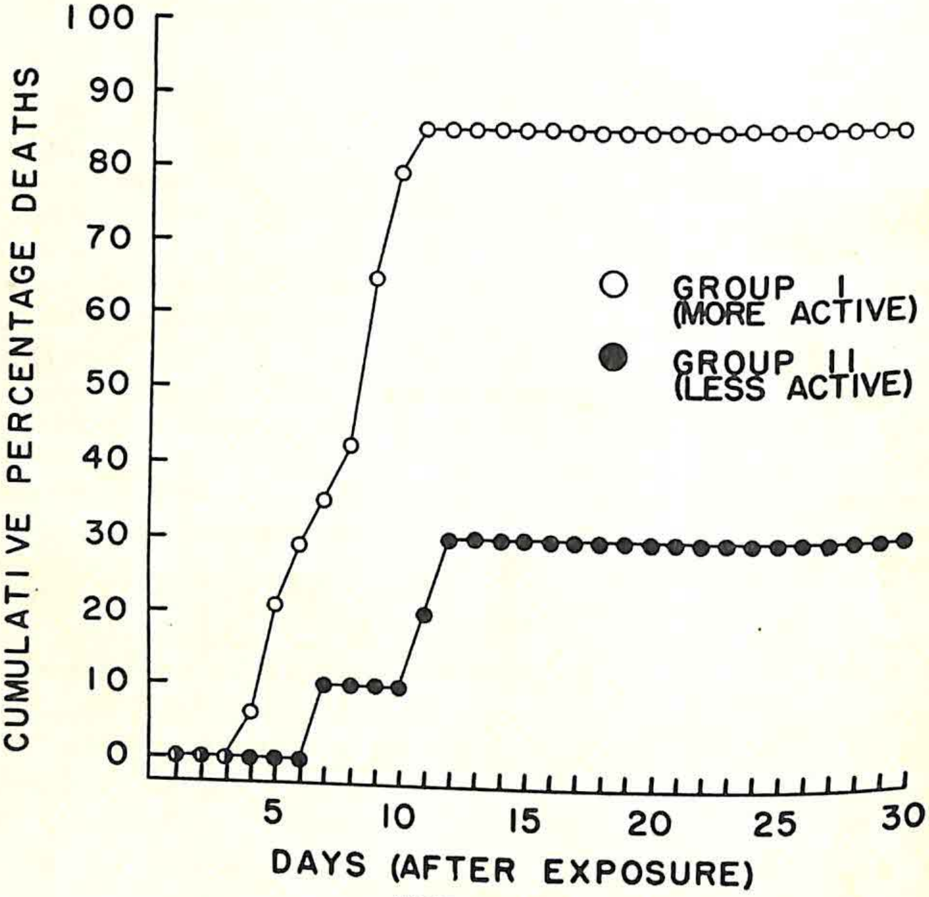

FIGURE 2
THE CUMULATIVE PERCENTAGE OF DEATHS IN EACH GROUP DURING THE FIRST 30 DAYS AFTER EXPOSURE TO RADIATION

pre-irradiation volitional activity. If this is the case, then the rat with the higher pre-irradiation rate of volitional activity would be forced by his individual rate to more exercise following irradiation than would the rat with the lower pre-irradiation rate, and hence, could be expected to die sooner, as demonstrated by Kimeldorf and Jones (3).

The findings of this study might be of considerable value to investigations of the physiological and biochemical bases of individual differences in susceptibility to irradiation. Rundquist and Bellis (4) have demonstrated that groups of rats differing significantly in activity level also differ significantly in metabolic rates. This suggests metabolic rate as a possible physiological hypothesis to account for differences in irradiation susceptibility—at least within the species.

E. SUMMARY

Mean and median activity rates (based upon revolutions per hour for hours during an eight-hour period in which revolutions occurred) for running wheel activity were determined for 14 young adult male Sprague-Dawley rats prior to exposure to 1200 r cobalt^{60}. Death order following radiation exposure was noted and correlations between both mean and median rates and death order were determined. Both measures were found to correlate significantly with death order, the animals with the higher rates dying first.

The study was replicated using 10 young adult male black non-agouti hooded rats. Again, both mean and median rates were determined to correlate significantly with death order. Also, the survival times of the animals of this group were significantly longer than those for Sprague-Dawley animals. This last phenomenon was predicted since the mean and median activity rates of the animals of the second group were significantly less than those of the animals of the first group.

REFERENCES

1. AUXIER, J. A. A high level Co^{60} irradiation facility. *Sch. Aviat. Med., USAF,* 1955, Report No. 55-40.
2. JONES, D. C., KIMELDORF, D. J., RUBADEAU, D. O., OSBORN, G. K., & CASTANERA, T. J. Effect of x-irradiation on performance of volitional activity by the adult male rat. *Amer. J. Physiol.,* 1954, **177**, 243-250.
3. KIMELDORF, D. J., & JONES, D. C. The relationship of radiation dose to lethality among exercised animals exposed to roentgen rays. *Amer. J. Physiol.,* 1951, **167**, 626-632.
4. RUNDQUIST, E. A., & BELLIS, C. J. Respiratory metabolism of active and inactive rats. *Amer. J. Physiol.,* 1933, **106**, 297-311.
5. SIEGEL, S. Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.

Department of Psychology
University of Texas
Austin, Texas

The Journal of Genetic Psychology, 1959, **94**, 217-220.

PARALLEL OBSERVATIONS ON AVOIDANCE BEHAVIOR AND ENDOCRINE RESPONSE UNDER RESERPINE*

Department of Psychology, University of Denver

STEPHEN B. COSLETT AND ALFRED B. SHAKLEE[1]

A. INTRODUCTION

The widespread use of tranquilizing drugs makes it of some importance to determine what effects such agents may have on simple dimensions of psychological response. Reserpine has been reported to decrease conditioned fear responses developed in an electric shock situation in rats (1) and monkeys (7). The amount of "protection" which reserpine may offer against the development of conditioned fear has in one case been reported as considerable (7), and in another as slight (6). The present research is concerned with the simultaneous effects of reserpine on shock avoidance behavior and on adrenal activation.

B. PROCEDURE

Female Wistar rats from 36 to 48 days of age at the time of initial injection were assigned at random to receive either 0.7 or 2.5 mg. per kg. body weight of reserpine or 0.2 ml. saline intraperitoneally on each of five consecutive days. The numbers of animals carried to the completion of the study were four, six, and six in the three groups respectively. Two animals were lost to the 0.7 mg. group due to apparatus difficulties. On the sixth day of their respective schedules animals were run in a shock avoidance shuttle to a criterion of five out of six correct responses or to a maximum of 75 trials. The shuttle box was patterned after the Mowrer-Miller (5) modification of Warner's apparatus, and had a grid floor, 5 in. by 24 in. The animals could learn to avoid a .2 ma. shock by moving to the opposite side of the apparatus during the sounding of a 2 sec. buzzer which immediately preceded current to the grid. The current, which was relatively independent of circuit resistance, was continued for a maximum of 22 sec. or until a correct response was made. Both intervals were controlled by a

---

* Received in the Editorial Office on February 28, 1958.

[1] The data relating to the shuttle response were collected by S.B.C., University of Denver, 1957, and tissue removals were performed by A.B.S. This investigation was supported by a research grant (M 1527) from the National Institute of Mental Health. The reserpine was kindly supplied by Dr. Jurg A. Schneider and Ciba Pharmaceutical Products, Inc.

synchronous motor timer. The time between trials varied between 1 and 2 min. Latencies were recorded from the onset of the buzzer to moment of correct response or to a maximum of 24 sec. Approximately 24 hrs. after the avoidance training adrenal glands and thymus were removed and weighed.

C. Results and Discussion

Body weight changes from the day of first injection to the day of avoidance training reflected overall losses for the two experimental groups (Table 1).

TABLE 1
EFFECTS OF RESERPINE ON AVOIDANCE BEHAVIOR, BODY WEIGHT, AND RELATIVE ADRENAL AND THYMUS WEIGHTS

	N	No. of animals reaching criterion*	Mdn. body wt. gain (g.)**	Mdn. adrenal wt. (mg.)*	Mdn. thymus wt. (mg.)**
Saline	6	5	+16	26.6	348
Reserpine (0.7 mg.)	4	0	− 4	25.8	266
Reserpine (2.5 mg.)	6	2	−17.5	23.1	170

* $p < .05$.
** $p < .01$.

Analysis of variance for weight changes in all groups by the H test yielded a value of 11.4, significant at the .01 level. Because of the small number of animals per group the two reserpine conditions were combined for analysis of error scores. Five out of six control animals reached criterion, whereas only two of the 10 receiving reserpine accomplished this. The difference has a two-tail probability of less than .05. Each animal's median latency was obtained for the first 11 trials—the number of trials in which the fastest learner mastered the habit. The overall comparison, yielding an H of 4.2 ($p > .10$ for 2 df), failed to show a significant difference on this index. In fact the central tendencies for latency were identical for the control and heavy dose reserpine groups, with medians of 3.1 sec. in each case.

Relative adrenal and thymus weights were computed in terms of mg. of gland per 100 g. initial body weight. Analysis of variance indicated that the experimental treatment produced both thymus atrophy ($H = 9.9$) and less significant but demonstrable adrenal atrophy ($H = 7.0$) (Table 1). The latter results are somewhat in contrast to those of Fluhmann (2), who found no adrenal weight change in young rats following 44–49 daily injections of 1.0 mg. per kg. body weight of an extract of *Rauwolfia serpentina*. Obvious differences in methodology could, of course, account for the discrepancy. The present outcomes may also be compared to the previous findings of Gaunt *et al.* (3), where dosages equal to or greater than 14 daily injections of 10 micrograms per 100 g. body weight produced adrenal hypertrophy and thymus atrophy in rats. Although the latter authors felt that there was no evidence for suppression of *ACTH* release by reserpine, possible leads to clarification of the demonstrated behavioral effects of reserpine may be found in pituitary effects like the weight reduction in this gland demonstrated by Fluhmann. The adrenal atrophy found in the present study could be a result either of depressed hypophyseal function or of a different mechanism which produced adrenal exhaustion. Mason and Brady (4) have shown in this connection that the suppression of conditioned anxiety in monkeys by reserpine administered over several days is paralleled by a depression of plasma 17-hydroxycorticosteroid levels. This decrease in blood corticoids following a series of injections of 0.75 mg. per kg. might be associated with an adrenal picture like that found in the present work. Whether the underlying mechanisms are similar in these two effects can only be established by further study. Possible hypophyseal involvement in reserpine produced behavioral changes will ultimately have to be tested against conceivable alternative mechanisms. Unfortunately, the loss of two animals and the high response variabilities precluded a good test of whether dosage level was related to magnitude of effects.

It is difficult to estimate to what extent the impairment in avoidance behavior may have been due to learning effects independent of motor processes. Although the interval of 24 hrs. between the last injection and the shuttle test was designed to allow for reasonable motor recovery, the large doses used could easily have left residual motor effects. While the groups did not differ significantly in terms of number of correct anticipations in 75 trials, the fact that six experimental animals had no more than two correct responses during training, while all controls had more anticipations, would make one especially cautious not to claim that the avoidance measure reflected learning. On the other hand, the similarity of the groups in terms of median latencies for the first 11 trials leaves the door open to the possibility that further research in this situation may demonstrate trials-to-criterion impairment not traceable to motor handicap.

D. SUMMARY

Five daily doses of reserpine were administered to female rats in amounts which were subsequently demonstrated to produce significant atrophy of the thymus and adrenals. This treatment also resulted in impaired avoidance behavior. The simultaneously demonstrated effects on behavioral and endo-

crine function offer one avenue for the physiological clarification of changes in certain instrumental response patterns.

References

1. Brady, J. V. Assessment of drug effects on emotional behavior. *Science,* 1956, **123**, 1033-1034.
2. Fluhmann, C. F. Effects of *Rauwolfia serpentina* on thyroid and pituitary glands of young rats. *Proc. Soc. Exp. Biol.,* 1956, **93**, 615-617.
3. Gaunt, R., *et al.* Endocrine aspects of the pharmacology of reserpine. *Ann. N. Y. Acad. Sci.,* 1954, **59**, 22-35.
4. Mason, J. W., & Brady, J. V. Plasma 17-hydroxycorticosteroid changes related to reserpine effects on emotional behavior. *Science,* 1956, **124**, 983-984.
5. Mowrer, O. H., & Miller, N. E. A multi-purpose learning-demonstration apparatus. *J. Exp. Psychol.,* 1942, **31**, 163-170.
6. Stein, L. Reserpine and the learning of fear. *Science,* 1956, **124**, 1082-1083.
7. Weiskrantz, L., & Wilson, W. A., Jr. Effect of reserpine on learning and performance. *Science,* 1956, **123**, 1116-1118.

Department of Psychology
University of Denver
Denver 10, Colorado

The Journal of Genetic Psychology, 1959, **94**, 221-232.

CONCEPTUAL BLOCK SORTING AS A FUNCTION OF NUMBER, PATTERN OF ASSIGNMENT, AND STRENGTH OF LABELING RESPONSES*

Department of Psychology, University of Massachusetts

HARVEY LACEY AND ALBERT E. GOSS[1]

A. INTRODUCTION

Fenn and Goss (3) hypothesized that learning common verbal responses to all stimuli within a particular conceptual category should increase intra-category similarity and that learning different verbal responses to stimuli in different conceptual categories should decrease inter-category similarity. In turn, these changes in intra-category and inter-category similarity were expected to facilitate the occurrence or acquisition of new sorting or choice responses to the same categories. To test this analysis of the rôle of mediating verbal responses in conceptual behavior, Fenn and Goss used normal and paranoid schizophrenic adults who, after learning discriminative verbal responses to blocks in four height-size categories, then sorted the blocks by height-size with markedly greater frequencies than their controls. In a related study with nursery school children, Carey and Goss (1) observed that prior strengthening of familiar word labels for blocks in height-size categories led to more rapid acquisition of placing the blocks by height-size. Goss and Moylan (4) then showed that frequency of height-size placements was a direct function of degree of mastery of different nonsense syllable or familiar word responses for each one of the four height-size categories into which the blocks could be placed. With a somewhat different task Hunter and Ranken (5) obtained transfer from color labeling to block sorting in terms of color.

The children of Dietze's (2) experiment learned highly similar or dissimilar names for particular combinations of shades of red and shapes of tips of 60 nonsense figures. In accordance with the hypothesis that dissimilar verbal responses should reduce inter-category generalization of choice responses, the dissimilar responses were associated with more rapid acquisition of the shade-shape concepts than were the similar names.

---

* Received in the Editorial Office on March 4, 1958.

[1] A more complete account is in (6). Parts of these data were presented at the meetings of the Eastern Psychological Association, New York, April, 1957.

B. Problem

The present study represents a further attempt to determine the affects of mediating verbal responses and stimuli on conceptual sorting behavior. The variables selected for investigation were: (*a*) the number of nonsense syllable responses for height and/or size categories, (*b*) the resultant patterns of assignment of the verbal responses to those categories, and (*c*) the degree of mastery of the responses.

Dietze varied the similarity of mediating response-produced stimuli and therefore, presumably, similarity of the shade-shape aspects of the stimuli by using sets of similar or dissimilar nonsense syllables. Another way of varying the similarity of the stimuli within and between categories is to use different numbers of verbal mediating responses. This procedure was followed in the present study in which four, three, or two nonsense syllables were assigned to blocks which could be categorized by height and/or size.

Five patterns of assignment were generated by these different numbers of labels for those categories. One pattern was assignment of a different nonsense syllable to each one of four height-size categories. The next two patterns involved assignment of different labels to two height-size categories and of a third syllable to the remaining blocks on the basis of height *or* size. In the last two patterns all tall or all large blocks were responded to with one syllable, and all short or all small blocks were labeled with a second syllable; thus assignment was on the basis of height *or* size.

Four labels, it was anticipated, would produce the greatest dissimilarity among the four height-size categories. Three labels were expected to increase the dissimilarity of two of the height-size categories; but responding to blocks within the remaining height-size categories with the same label should have the opposite effect of increasing the similarity of those blocks with respect to each other. Finally, there was the expectation that labeling blocks in two of the height-size categories with one syllable and using the other label for the other two height-size categories would further increase the similarity of the categories. Such assignment could only differentiate blocks in two of the height-size categories from the blocks in the other two height-size categories while it would increase the similarity of the blocks within each of the pairs of categories. Since transfer from verbal pretraining to sorting by height-size should decrease with increasing similarity of the stimuli within the height or size categories, it was predicted that extent of positive transfer to sorting by height-size would decrease as the number of labels decreased from four to three to two.

Fenn and Goss observed a stronger tendency to sort by height than by size. Should this tendency add to the effects of acquiring three or two labels for the categories, it seemed possible that learning three or two labels assigned partly or wholly by height might influence subsequent sorting by height-size somewhat differently than learning three or two responses assigned partly or wholly by size. The presumed direction of this influence was more height-size placements for assignment in part or wholly by height.

Transfer based on the number and pattern of assignment of verbal responses is also contingent on the strengths of those responses, which may interact with number and pattern of assignment. In order to assess such important interactions, *S*s learned the nonsense syllables to two degrees of mastery. Over-all number of labels-pattern of assignment combinations, the higher level of mastery was expected to produce greater positive transfer from verbal pretraining to block sorting than the lower level with, perhaps, interactions in the form of relatively greater effects of the higher level of mastery for increasing numbers of labels and for assignment by height rather than by size.

METHOD

1. *Subjects*

Classes in introductory psychology at the University of Massachusetts were the source of the 120 *S*s. They were randomly preassigned to 12 groups of 10 *S*s each.

2. *Stimuli and Apparatus*

The stimuli for both verbal pretraining and block sorting were 16 small blocks, each one of which was a different combination of tall or short, large or small and square or circular top and bottom areas, white or black (3). They were exposed one at a time by raising the shutter of the window of a boxlike exposure device; the shutter remained open until lowered by *E*. A 12-in. square wooden panel painted gray and divided by a ½-in. black line into four approximately 6 x 6 in. squares was added for the sorting task (3).

3. *Verbal Pretraining*

Table 1 shows the five patterns of assignment of four, three, and two labels to the height-size categories in combination with the two degrees of response mastery. Each of the four labels was assigned to a different height-size category; the abbreviation for this condition will be *4HS*. The third label of the three-label condition was assigned with respect to height *or* size. Since this also determined how the labels were assigned to the two height-size categories, there were different sets of blocks within the resultant *two* *height-size, one height* and *two height-size, one size* patterns of assignment. Henceforth, the former will be referred to as assignment of three labels by height (*3H*), and the latter as assignment of three labels by size (*3S*). Assignment of the two labels was in terms of *two height* (*2H*) or *two size* (*2S*) categories. The height and/or size categories of each of the five number-pattern of assignment combinations are noted beneath the names for those conditions.

TABLE 1
EXPERIMENTAL DESIGN

Number and assignment of labels	Degree of mastery 9/16	Degree of mastery 14/16
Four Height-Size		
Tall-large (*lag*), tall-small (*cev*), short-large (*mur*), short-small (*bik*)	10	10
Two Height-Size, One Height	10	10
Tall-large (*lag*), tall-small (*cev*), short (*mur* or *bik*)	(5)	(5)
Tall (*lag* or *cev*), short-large (*mur*), short-small (*bik*)	(5)	(5)
Two Height-Size, One Size	10	10
Tall-large (*lag*), short-large (*mur*), small (*cev* or *bik*)	(5)	(5)
Large (*lag* or *mur*), tall-small (*cev*), short-small (*bik*)	(5)	(5)
Two Height	10	10
Tall (*lag*), short (*bik*)	(5)	(5)
Tall (*cev*), short (*mur*)	(5)	(5)
Two Size	10	10
Large (*lag*), small (*bik*)	(5)	(5)
Large (*cev*), small (*mur*)	(5)	(5)
Instructions: No pre-sorting experience with blocks but instructed to sort by height-size	10	
Control: No pre-sorting experience with blocks	10	

The nonsense labels assigned to each category are in parentheses following that category. The two *3H* and the two *3S* groups were divided further into subgroups of five *S*s each to counterbalance any possible differential effects of which two of the four height-size categories were given different labels. Some categories of these two conditions are accompanied by two nonsense syllables in parentheses. One of these syllables was used with two *S*s and the other with three *S*s of each of the subgroups to counterbalance, at least partially, any possible differential influence of type of syllable. There was similar counterbalancing for each of the two-label groups.

Labels were learned by the paired-associates method to criteria of 9 or 14 correct anticipations in four successive units of four block-presentations (trials) each. These particular degrees of mastery were selected because previous findings (4) for the *four height-size* condition indicated that both degrees should result in positive transfer to block-sorting, but with significant superiority after mastery to the 14/16 level. As a further convenience in designation, level of mastery of particular label-pattern combinations will be indicated by following each of the abbreviations for number of labels-pattern of assignment combinations with 9/16 or 14/16. Because previous experimentation (1, 3, 4) had indicated that experiences in simply seeing and discriminating among the blocks had no effects on subsequent sorting, this prior discrimination training condition was not included.

The blocks were exposed one at a time with two-sec. anticipation intervals after which *E* told *S* the name; *S* then repeated the name. The order was random within the restriction that each one of the 16 blocks should appear within successive sequences of 16 blocks each. Inter-block intervals were not controlled but averaged about five sec.

The *instructions* condition was introduced to activate pre-experimentally learned labels for similar stimuli and experiences in sorting such stimuli into height-size categories. The effects on block sorting of this condition would serve as a reference point against which facilitation due to acquisition of discriminative labels through experimental training might be compared. Neither this group nor the *controls* saw the blocks prior to block sorting.

4. *Block Sorting*

There were four block-sorting trials during each of which 12 blocks were presented. At the beginning of each trial *E* placed four blocks in outside corners of the four squares of the sorting board. Each block represented one of the four height-size categories. A different set of four blocks was used on each of the four trials. Blocks representing a particular height-size category were placed in a different square on each trial. For each set of four blocks placed on the board, there were 12 remaining blocks, three within each height-size category. These blocks were presented one at a time under instruction to *S*s to place each one in the corner containing a matching block. In addition to the instructions given to the verbal learning and control groups, the *instructions* group was told that the blocks were to be placed by height-size; specifically, that they "should put all tall-larges, all tall-smalls, all short-larges, and all short-smalls together." No information was given to any group regarding correctness of placements nor was correction of incorrect placements permitted. Neither time allowed for placement nor inter-trial intervals was controlled.

Upon completion of the fourth trial *E* asked each *S* to tell him the principle or principles he had used for placing the blocks.

D. RESULTS

1. *Verbal Pretraining*

Means and standard deviations of trials to the 9/16 and 11/16 criteria are shown in Table 2. Means of the three and two label-pattern of assignment groups were first compared by means of a 2 x 2 x 2 analysis of variance (Table 3). The variables of this analysis were number of labels, pattern of assignment, and degree of mastery. The significant *F*'s of this analysis indicated that three-label groups took more trials to reach the same criterion than did two-label groups. As would be expected, the 9/16 criterion was reached in fewer trials than the 14/16 criterion. Finally, for both three-label and two-label conditions discriminative responses to differences in height were learned more rapidly than size discriminations.

TABLE 2
MEANS AND STANDARD DEVIATIONS OF TRIALS THROUGH VERBAL LEARNING CRITERIA AND OF HEIGHT-SIZE PLACEMENTS FOR TRIALS 1–4 OF BLOCK SORTING

Criteria	Labels	Trials through verbal learning criteria Mean	*SD*	Height-size placements for Trials 1–4 Mean	*SD*
9/16	*4HS*	56.0	23.0	21.0	9.8
14/16	*4HS*	73.4	20.2	34.7	17.2
9/16	*3S*	56.0	23.9	14.2	9.1
	3H	35.6	39.0	21.8	15.8
	3*	45.8	33.2	18.0	13.1
14/16	*3S*	75.2	32.1	22.2	12.1
	3H	47.6	25.7	29.2	13.7
	3*	61.4	31.7	25.8	13.1
9/16	*2S*	38.4	12.1	9.1	13.1
	2H	29.2	11.6	19.4	17.6
	2*	33.8	12.5	14.3	16.0
14/16	*2S*	48.0	18.4	15.8	14.4
	2H	36.2	20.1	10.4	6.4
	2*	42.1	19.7	13.1	11.2
Instructions		—	—	47.9	—
Control		—	—	11.6	9.5

* Combined height and size subgroups of three and two label conditions.

The four-label groups were included in the second analysis of variance (Table 3). The significant *F*'s for number of labels-pattern of assignment combinations and for degrees along with the nonsignificant interaction of these factors indicated that the 9/16 criterion was achieved in fewer trials than the 14/16 criterion. Also, for both criteria, the *4HS* and *3S* groups learned least rapidly at about the same rate, the *3S* and *2H* groups learned more rapidly at about the same rate, and the *2H* group acquired the syllables most rapidly.

TABLE 3

ANALYSIS OF VARIANCE OF TRIALS THROUGH VERBAL LEARNING CRITERIA FOR THREE AND TWO LABELS AND FOR FOUR, THREE, AND TWO LABELS

	Three and Two Labels			Four, Three and Two Labels		
Source	*df*	*ms*	*F*	*df*	*ms*	*F*
(A) Number of labels	1	4898.45	10.42**	4[a]	5316.47	9.26**
(B) Pattern of assignment	1	5951.25	12.66**			
(C) Criteria	1	2856.05	6.07*	1	4251.04	7.40**
A × B	1	911.25	1.94			
A × C	1	266.45	——	4[b]	3697.13	6.44**
B × C	1	120.05	——			
A × B × C	1	26.45	——			
error	92	470.20		90	574.33	

* Significant at $< .05$ level.

** Significant at $< .01$ level.

[a] Number of labels and patterns of assignment combined to give five number of labels-pattern of assignment conditions.

[b] Number of labels and patterns of assignment combined interacting with criteria.

2. *Block Sorting*

Means and standard deviations of numbers of placements by height-size for all four block-sorting trials are summarized in Table 4. All six of the four-label and three-label groups made more height-size placements than the controls. The two groups which learned two labels assigned by height also exceeded the controls. But the means of the *2S-9/16* and *2S-14/16* groups were slightly lower than that of the controls. In order to compare the means of the 10 verbal

TABLE 4

ANALYSIS OF VARIANCE OF CORRECT RESPONSES ON TRIALS 1–4 OF BLOCK SORTING

Source	*df*	*ms*	*F*
(C) Number of labels	2	374.50	8.11**
(B) Criteria	1	180.68	3.91
B × C	2	97.84	2.12
error (*b*)	94	46.18	
(A) Trials	3	17.51	9.46**
A × B	3	1.49	——
A × C	6	0.41	——
A × B × C	6	25.30	13.68**
error (*w*)	282	1.85	

** Significant at $< .01$ level.

learning groups with the mean for the control groups, variances of the former groups were averaged to give a value of 178.81. This average variance was then used to obtain the difference between the verbal learning groups and the control group necessary to yield a t significant at the .05 level for 18 degrees of freedom. The obtained value was 11.5. Considered singly, only the means of groups *4HS*-14/16 and *3H*-14/16 were significantly greater than the mean of the controls. However, eight or more of the 10 means of the verbal learning groups larger than the control group mean would occur by chance only six per cent of the time. Therefore, though most of the differences between individual verbal learning groups and the control group were not significant, the over-all pattern of the comparisons indicated that learning the nonsense syllables had facilitated sorting by height-size.

The mean number of correct placements by the *instructions* group was markedly higher than any of 10 means for the verbal labeling groups. Since this superiority would occur by chance only once in 1024 times, it can be concluded that knowledge of the height-size categories along with instructions to place the blocks in terms of those categories yielded more height-size placements than acquisition of discriminative naming responses.

Means of correct placements for three labels assigned by height and by size, and for two labels similarly assigned were compared by t-tests. For the *3H*-9/16 vs. *3S*-9/16 and *2H*-14/16 vs. *2S*-14/16 groups the obtained t's were 1.32 and 1.21, respectively. Neither these nor the t's of 1.48 for *2H*-9/16 vs. *2S*-9/16 and of 1.08 for *2H*-14/16 vs. *2S*-14/16 were significant. Data for assignment on the basis of height or size categories were pooled, therefore, for an analysis of variance (6) in which learning criteria and number of labels contributed to "between effects" and test trials constituted the "within factor." Number of labels, trials, and the interaction of these variables with criteria had effects significant at beyond the .01 level. The F of 3.91 for criteria was just short of the .05 level. However, the clearcut upward trend from controls to the 14/16 criterion for the four-label and three-label conditions leaves little doubt that height-size placement increased as a function of degree of mastery. If the control group had been included in the analysis of variance as a "zero" degree of mastery condition the interaction of criteria and labels would probably have been significant. The significant triple interaction is consistent with this suggestion.

The post-sorting query provided data on *S*s ability to state the height-size principle. An X^2 of 6.31 ($p = .05$ for 2 df) for 4, 3, 2 labels for both degrees of mastery combined indicated that the number of *S*s capable of stating the principle was directly related to the number of labels. However, the X^2 of

0.28 for the 9/16 and 14/16 learning criteria, disregarding number of labels, was not significant. For the 9/16 and 14/16 criteria combined, further significant X^2s for four labels ($X^2 = 12.94$; $p < .01$ for 1 *df*), three labels ($X^2 = 19.85$; $p < .01$ for 1 *df*), and two labels ($X^2 = 25.72$; $p < .01$ for 1 *df*) separately, and for the controls ($X^2 = 5.18$; $p = .05$ for 1 *df*), indicated that number of correct placements was directly related to the proportions of *S*s who were able to state the height-size principle.

E. Discussion

1. *Verbal Pretraining*

Both degree of mastery and number of labels were inversely related to trials to the verbal learning criteria. Though the evidence is indirect, the finding that in both three-label and two-label conditions height discriminations were learned more rapidly than size discriminations is consistent with Fenn and Goss' (3) observation that the pre-experimentally established preference for sorting by height is stronger than that for sorting by size. The groups learning three labels by height-size and size took as many trials to reach the 9/16 and 14/16 criteria as did four-label groups with corresponding criteria. Learning two labels by size was no faster than learning three labels by height. Thus number of labels and learning criteria interacted significantly.

2. *Block Sorting*

Comparison of all 10 verbal learning groups with the controls indicated that positive transfer to height-size placements had occurred. Four labels resulted in more correct placements than three, while two labels produced little if any advantage over the controls even when learned to the 14/16 criterion. The main effect of the 9/16 and 14/16 criteria of mastery was slightly short of statistical significance and the interaction of labels and criteria was also non-significant. However, if trends from the control or zero level of mastery to means for the four-label and three-label groups which had learned to the 14/16 criterion were considered, there is little doubt of a direct relationship between positive transfer to block-sorting and mastery of the labels. Numerical values for the 9/16 and 14/16 groups which learned four labels were similar to those reported by Goss and Moylan (4) for comparable *S*s. Also, the 14/16 value approximated those obtained by Fenn and Goss (3) for normal and paranoid adults. The interaction of number of labels and degree of mastery was not significant. However, for computational purposes, the means on which this interaction was based were restricted to those for the 9/16 and 14/16 degrees of mastery. Had the control group been included in the analysis as a zero degree of mastery condition the predicted interaction of a greater relative effect of number of labels for the 14/16 criterion would probably have been obtained.

No differences were obtained between groups discriminating by height or by size within particular combinations of patterns of assignment and degree of mastery. Thus differences in verbal learning rates for height and size discriminations with three labels and two labels alone, or possibly in combination with presumed initial tendencies to sort by height rather than by size, were not reflected in sorting by height-size.

Ability to verbalize the height-size principle increased with number of labels. In turn, there was a striking direct association between stating the principle correctly and number of height-size placements. This suggests that learning of the labels alone may not have been sufficient to produce extensive positive transfer to block-sorting. That is, if facilitation of block-sorting is to occur, such learning may have to be supplemented by ability to state the principle. If so, Fenn and Goss' (3) analysis of the bases of transfer from verbal learning to block-sorting may have to be extended to incorporate the rôle of knowledge of principles of block-response relationships. To do so, however, without greater knowledge of the nature and occurrence of such principles than is presently available does not seem feasible.

Possibly complementary to the hypothesis of increased intra-category similarity and decreased inter-category dissimilarity mediated by nonsense syllables is another interpretation. This view postulates that number of labels and/or degree of mastery increase the probability that *S*s will verbalize an already established height-size principle along with associated discriminative responses. Extent of facilitation of height-size placements should be a direct function of the degree of arousal of the principle and its related responses. Some of such self-instructed *S*s could not only be expected to state the height-size principle but also to place the blocks by height-size in the almost perfect fashion of those in the instructions condition. Whether introduced by *E* at the beginning of the sorting task or based on prior learning experiences, however, it is a plausible assumption that responding with the height-size principle increases the probabilities of appropriate discriminative responses to both blocks on the board and in the window. And blocks eliciting a common response should be placed together. Thus pre-established sets of discriminative responses which include the height-size principle and associated responses instead of nonsense syllables may serve to alter intra-category and inter-category similarity. The difference between the pre-experimentally acquired and the experimentally strengthened sets of responses, however, is in their origin and form. Both would presumably function to increase intra-category similarity and decrease inter-category similarity. Such pre-established sets of responses associated with the height-size principle are the assumed bases of the essentially errorless sorting by height-size of *S*s in the instructions condition.

F. SUMMARY

Conceptual block sorting was related to number, pattern of assignment, and degree of mastery of nonsense syllable responses to height and/or size dimensions of block stimuli. There were 16 blocks which represented combinations of tall or short, large or small and square or circular top and bottom areas, black or white. The four, three, or two labels assigned to the blocks by height and/or by size were learned by the paired-associates method to criteria of 9/16 or 14/16 correct anticipations. Each of the 10 groups which learned the labels consisted of 10 *S*s. Upon completion of verbal learning, four trials of 12 block presentations each were administered to test for transfer to sorting by height-size. Also given the block-sorting task were controls ($N = 10$) and an *instructions* group ($N = 10$), neither of which had had prior experience with the blocks. The *S*s of the latter group were told to sort the blocks by height-size.

Both number of labels and degree of mastery of the labels were directly related to number of height-size placements, but pattern of assignment by height or by size had no significant effects. The influence of number of labels was somewhat greater for the 14/16 criterion than for the lower level of mastery.

Analysis of *S*'s principles of sorting revealed that those who said they sorted by height-size made more height-size placements than those who could not state the height-size principle. Also, ability to state the height-size principle was directly related to the number of labels acquired.

One interpretation of these findings is that number and strengths of nonsense syllable responses influenced height-size placements by increasing intra-category similarity and decreasing inter-category similarity. Alternatively, the two experimental variables may have differentially activated the pre-experimentally acquired responses of stating the height-size principle and its associated discriminative verbal responses. Such sets of responses would also be expected to increase intra-category similarity and decrease inter-category similarity and thus influence subsequent sorting by height-size.

REFERENCES

1. CAREY, J. E., & GOSS, A. E. The rôle of verbal labeling in the conceptual sorting behavior of children. *J. Genet. Psychol.*, 1956, **90**, 69-74.
2. DIETZE, D. The facilitating effect of words on discrimination and generalization. *J. Exp. Psychol.*, 1955, **50**, 255-260.
3. FENN, J. D., & GOSS, A. E. The rôle of mediating verbal responses in the conceptual sorting behavior of normals and schizophrenics. *J. Genet. Psychol.*, 1956, **90**, 59-67.
4. GOSS, A. E., & MOYLAN, M. C. Conceptual block-sorting as a function of amount of discriminative learning. *J. Genet. Psychol.*, 1958, **93**, 191-198.
5. HUNTER, G. F., & RANKEN, H. B. Mediating effects of labeling on sorting behavior and judgments of similarity. Paper read at Eastern Psychological Association, Atlantic City, March, 1956.
6. LACEY, H. Conceptual block sorting as a function of type of assignment of verbal labels and strength of labeling responses. Unpublished M.S. thesis, Univ. Massachusetts, 1956.
7. LINDQUIST, E. F. Design and Analysis of Experiments in Psychology and Education. New York: Houghton-Mifflin, 1953.

Department of Psychology
University of Massachusetts
Amherst, Massachusetts

The Journal of Genetic Psychology, 1959, **94**, 233-251.

INFANTILE ILLNESS AND SUBSEQUENT MENTAL AND EMOTIONAL DEVELOPMENT*

Department of Psychology, University of Glasgow

D. H. Stott

A. Problem

The research into the causes of scholastic retardation conducted at the University of Bristol Institute of Education had as its central field-work a series of case-studies of a sample of backward children, in which a variety of data possibly relevant to mental and physical development were obtained. In successive reviews of these data the disproportionate amount of illness during the first five years of life, and in particular during the first two years, became increasingly striking. Among these illnesses pneumonia was especially frequent.

This association between scholastic retardation and serious early illness raised the question whether the latter was the primary cause, or was itself due to some other factor which was also responsible for the retardation. Among the possibilities that had to be formally tested was whether inferior social conditions could be responsible both for the retardation and the early ill-health. That there are important social factors in retardation has been demonstrated by Burt (4), and this was also apparent in the investigation in progress; but knowledge of a large number of 'good-family' cases suggested that the association between early illness and retardation held good independently of cultural level. Nevertheless it was necessary to confirm this formally, and to be in a position to examine other *prima facie* feasible explanations. If the influence of serious early illness were to be studied specifically the obvious step was to take a sample of children who had had such and compare their progress in school with matched controls.

A further interesting linkage was observed—that between backwardness and what was termed 'unforthcoming personality'. This is in essence an impairment of motivation, or more especially of the natural assertiveness and will-to-effectiveness which induces normal children to seek mental experience and to feel pleasure in overcoming difficulties. Children in whom motivation is poor lack confidence when faced with any new situation or problem even when it is within their mental or physical ability, are too timid

---

* Received in the Editorial Office on March 6, 1958.

to stand up for themselves among other children, and are 'mousey', shy and retiring in their general manner. In any group of educationally subnormal children these, it was found, formed the most prevalent personality-type, and it seemed to be mostly they who had suffered early illness. It was consequently thought possible that severe disturbance of the metabolism, very high temperature or some other unknown process might at a critical period inhibit the development of normal personality. The 'unforthcomingness' had indeed many points of similarity with the mood of lethargy or unresponsiveness to the challenges of life normal after severe illness in older persons, and it was wondered whether this could in fact be a fixation of such post-illness depression or its prolongation throughout the childhood. (Normal motivation was observed to return gradually, and in many cases to do so rapidly during puberty.) This suggestion made it necessary to secure personality-assessments of the early-illness group and of the controls.

B. Procedure

Facilities were granted[1] for a total sample to be obtained of all children who had spent at least two weeks as infants of 0–2 years in the Royal Hospital for Sick Children, Bristol, during 1946 and 1947, running into the first three months of 1948 for children admitted under one year. Children living outside the Bristol City Boundary were excluded, as also were illegitimate and foster-children and those who were found subsequently to have died. The net sample amounted to 198. The particulars of the diagnoses were obtained from the hospital records, together with those of later out-patient treatment if any.

At the beginning of the survey in the spring of 1955 the oldest child was about 11¼ and the youngest just over 7 years. The next step was to find which schools these children were attending. The School Welfare staff of the Bristol Education Committee were able to trace the majority, and in addition to make a general estimate of social category. Since it was essential for the validity of the survey that it should include as large a proportion of the original sample as possible, every effort was made to trace missing children. Resort to directories of residents was disappointing because it transpired that a large proportion of the parents of the untraced children had been living in 'rooms'. When, with the help of the Electoral Registration Office, such parents were identified and their movements followed, it became apparent that they constituted a shifting population moving from one part-house to another. In respect of the otherwise untraced more detailed enquiries were made at their old homes or at neighbouring houses, if they had not been bombed or pulled down. Where a neighbour could say to which other town or county the family had moved, the Education Authorities in question were asked to try to find them; all made the attempt, and when the children were found provided the required information as to school progress and personalities of the children. The final result of the search was that 142 of the 198, or 72 per cent, were found. Because the losses were largely among the non-householders, or other families who left no addresses with neighbours, the effect upon the sample was to eliminate a large part of the least prosperous and stable section. This accounts for the small proportion of ratings of 'inadequate standard' as given below.

---

[1] By Professor A. V. Neale, to whom, with Dr. T. Oppé of the Royal Hospital for Sick Children, Bristol, the author is also indebted for their guidance and interest in the present study.

Enquiry as to the children's progress had to be made at 72 schools, so that a refined assessment of 'reading age' would not have been feasible. Instead the head-teacher was asked to indicate whether the child read "better/about the average/not quite so well/is definitely backward/has hardly started to read at all." These were converted into a +1, 0, —2, —3 scale.

By way of a control-group the head-teacher was asked to choose a child of the same sex and of nearest birth-date whether in the same class or not. If there were two of equally near birth-date the one whose name was alphabetically higher was chosen. This procedure was not valid where the children happened to be in special schools for the educationally subnormal, and so the controls numbered only 133.

The personalities of both the hospital-children and the controls were assessed by means of the Bristol Social Adjustment Guides (3). These had the advantage of having been independently validated, being easy to complete and of providing a record of the child's behaviour and attitudes 'in the raw' and relatively free of the teacher's subjective interpretations. The items of behaviour were coded on to assessment forms according to their value as indicators of normality or of the maladjusted or other abnormal attitudes. The schedule offered the teacher, for marking in respect of each child, some 300 items, so that a fairly detailed sounding was obtained of the social adjustment.

Apart from the testing of the hypothesis that impairment of motivation was the causal link between early illness and retardation it was anticipated that the assessments of personality would throw fresh light upon the effects of hospitalisation as such and at various stages of the infancy.

The main purpose of the enquiry, however, was to test the connection between early illness and retardation, and if this was confirmed, to give an indication of the extent to which children in general who suffer illness in early life are likely to be scholastically retarded. The teachers' assessments of reading ability are summarised in Table I.

TABLE 1
TEACHERS' ASSESSMENTS OF READING ABILITY

	Hospital children (142)	Controls (133)
Above average (+1)	19	22
Average (0)	45	67
Slightly below (—1)	39	29
Definitely backward (1–2)	30	10
Non-reader (—3)	9	5

C. RESULTS

It will first be noted that the controls, with only 22 children above the average and 44 below, are shown as of somewhat below average ability. This is open to various explanations. Studies of the distribution of intelligence-test scores regularly show a hump at the lower end, so that a symmetrical distribution of reading ability is not to be expected. Probably, however, the main reason is that teachers of primary-school children tend to set their standards too high, and so to rate children down. Since they would have used the same standard for the hospital-children as for the controls, this bias does not affect the comparison. However, because some of the 'Slightly below average (—1)' group would have been in fact of average ability, this class is best treated where the number of cases is sufficient as an indeterminate borderline group. This gives a 6-cell arrangement in which the distribution of the —1's, while showing the same general tendency as that of the more extreme backwards, is not far off the expected chance proportions, and the value of χ^2 is made up chiefly by the disproportionate incidence of definite backwardness among the hospital-children. With $\chi^2 = 15.0$ and $n = 2$ this is significant at less than the 1000 to 1 risk of chance.

Table 2 shows the arrangement of the above figures upon which the calculation is based, with the percentage of children in each category.

Even without a statistical test, the early-hospital children are seen as a group to be significantly more retarded in reading than the controls (an unknown proportion of whom would also have been in hospital as infants). Nevertheless 45 per cent of the hospital-children were of average reading attainment or above, and very nearly as great a proportion—13.4 per cent compared with 16.5 per cent in the controls—were of above average ability.

This suggested more detailed enquiry as to which of those children who had been in hospital as infants were the most severely retarded.

TABLE 2

	Reading attainment			
	−3, −2	−1	0, +1	
Hospital-Children	39 (27.5%)	39 (27.5%)	64 (45%)	142
Controls	15 (11.3%)	29 (21.8%)	89 (67%)	133
	54	68	153	

In order to assess the influence of the cultural level, the families of the hospital-children were placed in the five categories indicated in Table 3, which were standard throughout the research.

TABLE 3

I	Middle class, professional	8
II	Lower middle class, skilled worker with similar standard	19
III	'Respectable' working class, clean but unpretentious	96
IV	Low-culture, 'labourer' type	10
V	Slum, very dirty, 'problem-family'	—
	Not known	9
		142

As explained above, the small number of families of inadequate standard is accounted for by the loss of the untraced 'tail' of those who at the time of the children's admittance to hospital were not living in houses of their own. For this reason the hospital-children were probably of a somewhat higher average social level than the controls (of whom it was not possible to make any assessment of social standard).

The general association between cultural level and reading ability in the early-hospital group is shown in Table 4.

TABLE 4

	Cultural level	−3, −2	−1	0	+1	
I.	Middle class	—	—	2	6	8
II.	Lower m-c.	3	3	10	3	19
III.	General working class	29	30	28	9	96
IV.	Low-standard	6	2	2	—	10
V.	Slum	—	—	—	—	—
	Not known	1	4	3	1	9

It is seen that there are no backward children in the middle-class, and very few among the lower-middle group; conversely there are no forward children in the low-standard group, the majority of those in the latter being in fact in the 'definitely backward' or 'non-reading' categories. Whatever effect early ill-health and/or hospitalisation has, therefore, upon subsequent reading ability, or whatever other factor is common to both, is strongly conditioned by social standard. Since the great majority of the children were members of clean but socially unpretentious working-class families, the easiest way of ascertaining whether an association between early ill-health and retardation exists independently of social class is to treat the large Group III as a sample on its own and compare it with the controls. This is shown in Table 5.

TABLE 5

	Reading standard −3, −2	 −1	 0, +1	
Hospital children (social Grade III only)	29 (30.2%)	30 (31.3%)	37 (38.5%)	96
All controls	15 (11.3%)	29 (21.8%)	89 (67.0%)	133

$\chi^2 = 20.5$, $n = 2$, $p. < .001$.

It happens that by the removal of both the inferior and superior grades the proportion of backwardness among the hospital-children is actually increased, and the risk of a chance result falls well below 1000 to 1. There is consequently seen to be an association between early ill-health and later scholastic retardation independently of social class, although it cannot be assumed that the connection is one of cause and effect.

The next step is to enquire whether any particular illness or category thereof can be picked out as being specially closely associated with later retardation. The disorders for which the children were treated while in hospital as infants were divided into the categories as shown in Table 6, beside which the incidence of later retardation is given. Since this classification is of illnesses, the total is greater than the number of children.

TABLE 6

	Reading Attainment −3, −2	 −1	 0	 +	
Number of children	39	39	45	19	142
Pneumonia	12	5	3	1	21
Other respiratory	5	7	6	2	20
Digestive system	10	13	12	4	39
Local infection	9	7	14	5	35
Other illness or defect	18	20	13	7	58

The one category of illness which stands out as closely associated with retardation is pneumonia. Over half the children diagnosed as having had this illness as infants were either definitely backward or completely non-readers, and only four were of average reading attainment or above. Two were mentally deficient. Only four of the 21 pneumonia cases were in the inadequate cultural grades, and three of these were retarded, so that cultural influences could not account for the retardation of the pneumonia-group as a whole.

No other class of illness marks itself out in this way, but in all there is a consistent skewness of the distribution towards the 'retarded' side. If the children who had had pneumonia are removed the remainder are still seen in Table 7 as more retarded than the controls.

TABLE 7

	−3, −2	−1	0, +1	
Hospital children less pneumonia cases	27	34	60	121
Controls	15	29	89	133

$\chi^2 = 9.0,\ n = 2,\ P \equiv .01.$

Even excluding the pneumonia cases the association between early-illness and scholastic retardation remains significant at very nearly the 100 to 1 risk of chance.

A noticeable feature was the large number of infants who were diagnosed as suffering from two or more apparently quite separate complaints. In view of a similar tendency to multiple early illness found in another sample of children in respect of whom stresses during the pregnancy were reported (15), the present sample was divided into two groups—those treated in hospital for one disorder only, and those treated for two or more different disorders under the age of two years. The effect might be to divide them very roughly into two broad classes, namely those in which a considerable proportion of the illnesses were due to high exposure to infection or to poor maternal care, but who were otherwise healthy children; and those who had some more general physical impairment or susceptibility which showed itself in apparently unrelated disorders or defects. The latter category might be arbitrarily termed the "weak-constitution" group (Table 8).

TABLE 8*

	Backward −3, −2, −1		Not backward 0, +1		% Backward
One type of illness only	22	26	54	102	47
Two or more types of illness	15	12	10	37	73

---

* Two mentally defective children were in hospital for observation only, with no recorded illness, and in one other case the type of illness was unknown. This reduces the total of the above to 139.

This shows a preponderance of backwardness among those children treated for more than one disorder—73 per cent as opposed to 47 per cent—which gives a χ^2 of 7.26 ($p < .01$ if the —1's are included among the backwards and nearly .02 if they are treated as a separate category).

The great majority of the disorders fell into three main groups: respiratory, digestive, and other localized infection. The fourth, miscellaneous group, were in large part congenital malformations. Illnesses were only counted as of different types for the purpose of defining 'weak constitution' if they fell in different categories as detailed in Table 9, except that two illnesses of Category 4 were treated as different if they related to different parts of the body and did not belong to a known syndrome.

TABLE 9

Categories of Illness and Physical Defect Treated as of Different Types

Respiratory:

- Pneumonia
- "Acute respiratory infection"
- Acute bronchitis
- Cough, recurrent colds, (and variants of above wording)
- Pulmonary tuberculosis

Digestive:

- Enteritis
- Pyloric stenosis
- Kidney disease
- Coeliac disease
- Appendicitis
- Intussusception
- Diarrhoea & vomiting
- Feeding difficulties
- Constipation
- Duodenal obstruction

Localized infections:

- Otitis media
- Discharging ears, mastoid
- Inflamed throat, tonsilitis
- Discharging eyes
- Abscesses
- Skin infection
- Streptococcal infections
- Enlarged glands
- Septic arthritis
- Non-pulmonary tuberculosis

Other illnesses:

- Epispadias
- Urticaria
- Meningitis
- Encephalitis
- Poliomyelitis
- Syphilis
- Macrotic Anaemia
- Pink Disease
- Cardiac disease
- Odoema
- Rectal Prolapse
- Hernia
- Cleft Palate, hare lip
- Naevus
- Circumcision
- Hydrocephalus
- Incision—gland of groin
- Cavernous Laematoma of scalp
- Fontanelle defect
- Hallux Varus
- P.U.O.
- Fits, Convulsions

Any explanation of these associations must at the present stage be tentative. Whereas the high risk of subsequent retardation following pneumonia suggests some specific damage to the central nervous system arising from high temperature or other cause, the factor of multiplicity points to a more general impairment of which proneness to diverse diseases and possibly mental subnormality may be a part. This latter is consistent with the hypothesis of pregnancy/general-impairment (*P/GI*) arrived at in the study of another group of mentally defective and subnormal children mentioned above (15). Close comparison of the data from the two samples suggests in fact that both etiologies may be operative. Among the 40 mentally defective children of this earlier sample there were three whose retardation dated from pneumonia under three years of age, with fairly good evidence of prior normal development, while the retardation of four other pneumonia cases seemed to be congenital. These latter cases and others of acute respiratory disease were associated with pregnancy-stress. Consequently the pneumonia cases among the present sample of early-hospital children may have a dual etiology, one group being due primarily to strong exposure to infection of robust children, the other being due to constitutional susceptibility following an unfavourable pregnancy.

It cannot be assumed that all the cases with one illness only were robust children who had suffered high exposure to infection, neglect, or injury. Among the mentally defective and subnormal of the earlier samples who were ill in infancy subsequent to pregnancy-stress, 49 per cent had more than one serious illness—a rather greater proportion than in the present study. If therefore, pregnancy-stress is the common factor underlying both the proneness to disease and the retardation in the present sample, this would also apply to a proportion of those who up to their hospitalisation had only had one disease. Without very detailed data as to the condition of the children from birth it is naturally difficult to make an estimate of the proportions who were disease-prone under conditions of normal exposure, and those who were not so but who were abnormally exposed. The criterion of multiplicity of disease is thus a very rough one, and for this reason its significant association with backwardness would suggest the existence of some very powerful underlying factor.

Of all the childish complaints, those of the digestive system are probably most liable to be caused by abnormal exposure to infection. It is consequently useful to apply to them the criterion of conjunction with other disease (Table 10).

It is seen that nearly all the not-backward children fall in the group who had a disease of the digestive system and no other. Despite the small numbers this is significant at 100 to 1 risk of chance ($\chi^2 = 6.73$). The same tendency was apparent in the cases of respiratory infection excluding pneumonia, but the numbers were too low to be of significance.

TABLE 10

	Backward (−3, −2, −1)	Not backward (0, +1)		% Backward
Disease of digestive system only	9	13	22	41%
Disease of digestive system with other disease	14	3	17	82%
	23	16	39	

It has been mentioned that the findings of the main study had in the meantime suggested a rival hypothesis to account for the connection between early illness and subsequent retardation—that of pregnancy-stress as the common underlying factor. The original suggestion had been that the acute early illness inflicted injury upon the nervous system, either by causing some impairment to the cognitive capacity of the brain, or by an impairment of motivation which would result in non-use or poor use of whatever capacities existed. The hypotheses are not exclusive, and with pneumonia there is some evidence that both forms of impairment can operate. Nevertheless, with the strong connection that has been demonstrated, it is interesting to enquire what further evidence supports one hypothesis or the other as the main explanation. The causal link in the original theory was that of impairment of motivation as a result of the early illness. This can be tested by observing the association between such motivational impairment and backwardness on the one hand, and the types and number of illnesses on the other.

TABLE 11

READING ABILITY AND PERSONALITY IN EARLY-HOSPITAL CHILDREN*

	Normal personality	Slight unforthcomingness	Some/strong unforthcomingness UF	Unsettled & maladj. with UF	Unsettled & maladj. without UF	Total
Reading						
−3, −2	5	4	20	1	4	34
−1	7	3	17	1	10	39
0, +	33	3	18		10	64
	45	10	55	2 +	24	137

* Four mentally defective children excluded as criteria of social adjustment cannot be rigidly applied, and in one case the social adjustment could not be ascertained.

Table 11 shows the relationship between reading ability and impairment of temperament or personality. The assessment of the latter was based on the standard categories used for previous studies, but since there were comparatively few maladjusted children of the actively anti-social or chronically restless and disorganized types, these were grouped under one heading of 'Unsettled and Maladjusted'. This left, apart from the children of stable temperament, three grades of an inhibited or unforthcoming behaviour-pattern of varying severity (depression and withdrawal as well as the 'pure-type' unforthcomingness being grouped together for the purpose of this study). Earlier studies had shown that the slight manifestation of unforthcoming behaviour—not more than four items reported—was apparently cultural in origin, and not much related to backwardness or early illness. It was therefore here grouped as a variant of normal behaviour, that is to say, merely as a certain reticence or shyness. In any case the small numbers and their neutral distribution hardly affect the significance. One further small complication was that two of the unsettled or maladjusted group also qualified for a rating as unforthcoming, and so had to be treated as such; they have a sub-column to themselves.

The first point to be noted is the comparatively small amount of unsettledness and maladjustment in these children who had spent at least two weeks in hospital as infants, actually a smaller proportion than among the random controls (17.7 per cent as opposed to 24.6 per cent). The excess in the controls is probably due to the smaller field among the hospital-children for the more active type of unsettledness once the large group of the 'unforthcomings' are subtracted. Of those who qualified for a definitely 'maladjusted' rating, there were nine among the 137 hospital-children and eight among the 130 controls for whom social-adjustment ratings were obtained. There is thus no indication that early hospitalisation of short duration—or from two to four weeks for the most part—has any permanent effect in producing maladjustment.

By far the most striking feature of the above figures, however, is the heavy preponderance of stable children among the average or better readers, which is compensated by the disproportion of unforthcoming children among the poor readers. The comparatively strong relationship between unforthcoming personality and poor reading is shown in Table 12.

TABLE 12

	Reading attainment		
All hospital children	−3, −2, −1	0, +1	
Unforthcoming	39	18	57
Not unforthcoming (except in normal slight degree)	34	46	80
	73	64	137

$\chi^2 = 9.12$, $p. < .01$, nearly .001.

It is seen that over half the backward readers have unforthcoming personalities, as opposed to only 28 per cent of the average or good readers. This association can in a small degree be accounted for by cultural factors. In the lowest social group represented, that of the low-culture labourer-type family, all the unforthcoming children were backward, but in the highest (middle-class professional) group the three unforthcoming children were all average or forward readers. This suggests that the lack of confidence and of normal assertion which is the central feature of the unforthcoming temperament is a definite handicap to reading progress at the low cultural level, but that in an education-conscious social group means are found of overcoming such children's fear of facing difficulties. The lower-middle class group occupies an intermediate position, with five of the unforthcoming children being backward and three not. Nevertheless this interesting relationship between motivation and attainment remains very nearly as close among the ordinary working class bulk of the sample. Table 13 gives comparable figures as in the previous table for this social group only.

TABLE 13

Hospital children of social grade III only	Reading attainment $-3, -2, -1$	 $0, +$	
Unforthcoming	28	11	39
Not unforthcoming (except in normal slight degree)	27	26	53
	55	37	

$\chi^2 = 4.29$, $p. < .05$.

The unforthcoming group still account for just over half the backward, but only 30 per cent of the adequate readers. The lower χ^2 is mainly due to the smaller numbers. There would consequently seem to be some fundamental relationship between unforthcoming temperament and backwardness. It is found also in the control group ($\chi^2 = 4.18$), and indeed this personality-type is likely to be the most prevalent one in any group of educationally subnormal children. If the motivational factor is primary, in the sense that the lack of confidence, normal assertion, and curiosity, etc., entails a limitation of mental experience and hence of concept-formation, this could mean that the retardation was caused more by a temperamental than a cognitive impairment. That is to say, the poor attainment and probable poor general mental development may often be due more to lack of use of the brain than its actual damage, poor quality, or congenital impairment.

Thus far the evidence quoted still remains open to two interpretations.

The early illness, by pyrexia or otherwise, could impair motivation—thus affecting school attainment—or both the early illness and the defective motivation could be joint effects of a more fundamental cause in the pregnancy, which in addition was liable to produce mental impairment. The latter hypothesis implies that the elements of impairment following a disturbed pregnancy are sporadic, as indeed the main study suggested. There it was found that impairment of health in infancy, mental retardation of an undoubtedly congenital type, and congenital malformation were all independently associated with pregnancy-stress, but in varying proportions.

The present study poses two questions in relation to this hypothesis of pregnancy/general-impairment: the first is whether temperamental impairment is also an element of this pregnancy-stress syndrome; the second is whether mental retardation subsequent to early illness is found independently of temperamental impairment in its most common form, namely unforthcomingness (18). The first can only be approached indirectly because no pregnancy-histories for this group of hospital-children were available.

It has already been seen that the children who had two or more disorders were significantly more backward than those who had only one. Since the types of illness or defect were apparently distinct, and moreover a considerable proportion of those in the miscellaneous fourth category were congenital malformations, this suggested some underlying congenital common cause. There was about the same incidence of unforthcomingness among the double-illness as among the single-illness children, which is a slight indication against its being part of the *P/GI* syndrome. On the other hand unforthcomingness cannot be explained solely as an after-effect of pyrexia or other acute illness, because in fact it was also prevalent among those single-illness cases in which there was no question of such. There were 41 unforthcoming children who were treated for only a single disorder, and the latter can be broken up into two categories: the first consisting of those which by pyrexia or secondary meningitis could conceivably cause neural damage resulting in motivational impairment, and those in which this is unlikely or inconceivable (Table 14).

It is seen that only a minority of the cases of impairment of motivation can be feasibly attributed to the illness itself; in certain instances, notably the congenital malformations, abscesses and other localized infections, such sequelae are hardly conceivable. This suggests that the unforthcomingness is an independent symptom which may be congenital, and so may also be the result of pregnancy-stress.

In view of the well-known contemporary controversy (1, 2, 5, 6, 7, 8, 9, 10, 13) the possibility must be considered that the impairment of temperament was the result of the early hospitalization. It must first be emphasised that although the unforthcoming personality-type may very broadly be described as inhibited or even withdrawn, it does not in the least conform to Bowlby's concept of 'affectionless character'. Only two of the 57 'unforthcomings' showed signs of hostile or other anti-social maladjustment, and the unforthcoming behaviour of these, although it was rated as such by the objective scoring criteria used, was more probably due to the depression characteristic of the hostile-anxious maladjusted child. Unforthcoming children are generally over-affectionate and dependent, and loveable good children whose capacity for hostility is lower than the normal.

TABLE 14
SINGLE-DISORDER UNFORTHCOMING CHILDREN

I		II	
Respiratory:—		Digestive:—	
pneumonia	7	pyloric stenosis	2
other	6	appendicitis and intussusception	2
Wasting, not thriving	2	coeliac disease	1
Meningitis, encephalitis	2	vomiting and diarrhoea	1
Haemolytic strep. tonsilitis	1	Urticaria	1
		Abscesses, cellutlitis	7
		Otitis media	2
		Congenital malformation apart from any above:	
		naevi	2
		cleft lip/palate	4
		hernia	1
Disorders possibly causing neural lesion and notably impairment of temperament	18	Disorders unlikely to cause such	23

At the stage in the study of the main group of backward children when the phenomenon of unforthcomingness was attracting interest, an analysis was made of a possible connection between such and early separation (14). From the 141 case-histories then to hand, 25 were chosen who had been separated from their mothers for at least 10 weeks during their first four years of life. This minimum was however far exceeded for the majority of the group, the average length of separation during the first year alone being 18 weeks, and for the second year 17 weeks. Of the 25, eight were well-adjusted personalities with normal motivation, eight were unforthcoming, and nine were unsettled or maladjusted. Of the 28 children within the same sample of 141 who were seriously ill during their first two years but not hospitalized or otherwise separated from their mothers, only five were well adjusted, 15 were unforthcoming, and eight unsettled or maladjusted. There were, then, actually more unforthcoming children among those who had been ill at home with their mothers than among those who had been ill in hospital. This is not significant owing to the small numbers, but it suggests that unforthcomingness is more closely connected with serious early illness than with being in hospital.

Within this '141' sample of backward children there were 51 with unforthcoming personalities; in Table 15 the relationship of such with *any* separation and with early illness respectively is analysed.

TABLE 15
UNFORTHCOMING BACKWARD CHILDREN IN MAIN STUDY

	Separated from mother at least one week 0–2 yrs.	Not separated	
Serious illness or equivalent 0–2 yrs.	18	15	33
Not seriously ill	2	16	18
	20	31	51

It is seen that 33 of the 51 unforthcoming children had been seriously ill in infancy, and only 20 separated; and of the latter only two had been separated without being ill. This established a definite connection between the unforthcomingness and serious illness, and a striking lack of association between this form of temperamental impairment and early separation from the mother. With $\chi^2 = 9.54$ this gives not far short of a 1000 to 1 risk of chance, and can therefore be regarded as conclusive.

This does not imply nevertheless that the unforthcomingness can be regarded as a consequence of the illness, for there were 18 of the 51 unforthcoming children who had not been seriously ill. The distribution as a whole is however consistent with the hypothesis of a multiplicity of impairments following a disturbed pregnancy. If this is so, the unforthcomingness (or impairment of temperament) would be one of the elements of a syndrome of general impairment which includes early ill-health and mental retardation. There is, therefore, some evidence in favour of the second hypothesis under consideration, that the illnesses and the temperamental defects of the early-hospital group of children which are the subject of the present paper were in some degree common results of factors operating within the pregnancy.

By this hypothesis the temperamental impairment in the form of unforthcomingness would be expected to manifest itself sporadically in the same way as the other elements of the *P/GI* syndrome. From this it follows that even when it is absent, the remaining elements of the syndrome should still be found to be associated with each other. We come now consequently to the most crucial test of whether the backwardness was produced by the early illness *via* impairment of motivation, or whether it was at least sometimes due to congenital mental impairment as a factor in itself apart from unforthcomingness. In other words, it is a question of testing the association between backwardness and 'weak constitution' (multiple disorder) in the not-unforthcoming children only (Table 16).

TABLE 16
NOT-UNFORTHCOMING CHILDREN ONLY

	Reading attainment			% Backward
	−3, −2, −1	0, +1		
One disorder only	21	38	59	35.5
Two or more disorders	13	8	21	62.0

It is seen at once that the multiple-illness children, even without unforthcoming personality, tend to be more backward than the single-illness children; although the numbers concerned are small the χ^2 of 4.46 is significant at less than a 1 in 20 risk of chance. This weighs in favour of the hypothesis that part of the mental retardation is congenital. Once more it is seen that the majority (62 per cent) of the multiple-disorder children are backward. On the other hand the proportion of backwardness is somewhat lower (see Table 8) than when the unforthcoming group are included. This is open to two explanations. The first, on the pregnancy/general-impairment hypothesis, would be that the appearance of one element of the syndrome, here mental retardation, would be more likely if two others were already present (illness + unforthcomingness) than if only one (namely illness) were; since illness alone may be unconnected with the syndrome. The second is that the backwardness, as an aspect of poor mental development, is partly the result of poor motivation. There is in effect no reason why both causal linkages—the one directly connecting retardation with congenital impairment, and the second showing it as the corollary of congenitally defective motivation—should not be feasible, with one applying to one case and the other to another, or operating in combination. Clinical experience of retarded children suggests that those of the educationally subnormal type—of *IQ* 50–80—tend to be handicapped more by impairment of motivation, while below this level direct cognitive (mostly association—and concept-forming) impairment seems evident.

As a final means of assessing these two etiological possibilities the same type of analysis as that made in respect of unforthcomingness in Table 14 can be repeated for backwardness. The disorders are grouped in (I) those which, by pyrexia, cerebral inflammation or otherwise, may be held likely to cause direct neural lesion, and (II) those unlikely to do so (Table 17).

TABLE 17
SINGLE-ILLNESS CHILDREN ONLY

	Backward (−3, −2, −1)	Not backward (0, +1)
I. *Disorders possibly causing cerebral lesion:—*		
Respiratory:—		
pneumonia	11	3
T.B.	3	2
other	5	5
Wasting, not thriving	—	3
Meningitis, encephalitis	4	2
Syphilis	—	1
Poliomyelitis	—	1
Cavernous Laematoma	—	1
Macrotic anaemia	—	1
	23	19
II. *Disorders unlikely to cause cerebral lesion:—*		
Of digestive system	9	11
Urticaria, pink disease, epispadias	1	2
Abscesses, E.N.T., enlarged glands	9	15
Congenital malformation apart from any above, naevi, hernia, rectal prolapse, circumcision	6	7
	25	35

Such a division of illnesses based on short notes of diagnosis made some years back must of course be inaccurate. For example the "Wasting, not thriving" group of illnesses might equally well have been in either category; but to be on the safe side an illness was put in Category I when there was any doubt, as the object at this stage of the analysis was to examine the subsequent attainment of the children whose early illnesses could not be regarded as directly responsible for their backwardness. It is seen that nearly half the backward children were in Category I, compared with 35 per cent of the not-backward children. This difference is not significant, which suggests that while the backwardness was not attributable to the sole factor of direct cerebral lesion from the disease itself, the latter cause was probably operative in a proportion of the cases. This can once more be interpreted as evidence that the retardation of the hospital-children was of dual origin.

D. SUMMARY

Following the observation of a very high incidence of infantile illness in the case-histories of backward children, a study was made of the school progress and emotional development of all the legitimate, Bristol-born children who had spent at least two weeks in a hospital between 7 and 11 years ago. Seventy-two per cent of these were traced to their present schools, and were matched against controls whose previous history was unknown.

The children known to have been in hospital were significantly more retarded: 27.5 per cent were non-readers or definitely backward compared with 11.3 per cent of the controls.

The cultural level of the families proved to be a factor in scholastic success, but this did not account for the backwardness of the early-hospital children. Among those from ordinary working class families with adequate standards of care and hygiene 30.2 per cent were backward.

The causal link was sought in the illnesses themselves. The most striking revelation was that 12 of the 21 early-pneumonia cases were definitely backward. Nevertheless the 121 non-pneumonia cases were still significantly more backward than the controls.

In view of the tendency, observed in a previous study, to multiple early illness subsequent to stress during the pregnancy, the sample was divided into those who had only one type of illness or defect, and those who had two or more distinct types: 73 per cent of the latter group were backward in some degree, compared with 47 per cent of the single-illness group, which was again highly significant. This suggested that there might be some general factor acting as common cause of 'weak constitution' and of mental impairment.

A significant proportion of the backward children were of 'unforthcoming' personality as assessed by the Bristol Social Adjustment Guides. Since this consists basically in impairment of motivation, the suggestion was made that it might provide the link between the early illness and the backwardness. This raised the question whether this impairment of personality was due to the early illness or was itself part of a syndrome of general impairment following stressful pregnancy.

It was found that only a minority of the 'unforthcoming' children had suffered from infantile diseases likely to produce cerebral lesion or other permanent after-effects. Moreover there was still a significantly higher proportion of backwardness among the multiple-illness children within the class of those who suffered no impairment of motivation. While the latter could not be ruled out as the causal link in some cases, the evidence pointed to a common congenital origin for the early illness, the mental retardation and the impairment of personality termed 'unforthcomingness'.

In an earlier study of backward children a similar association was found between unforthcomingness and early illness; but this form of personality-impairment was quite unrelated to separation from the mother in infancy. Thus the short stay in hospital—mostly of 2 to 4 weeks—could be ruled out in the present study as responsible for subsequent emotional, and presumably mental, impairment.

References

1. Bowlby, J. Forty-four Juvenile Thieves. London: Bailliere, Tindall & Cox, 1947.
2. ———. Maternal Care and Mental Health. Geneva: World Health Org., 1951.
3. Bristol Social Adjustment Guides and (manual) the Social Adjustment of Children. London: Univ. London Press, 1958.
4. Burt, C. L. The Backward Child. London: Univ. London Press (3rd ed.), 1950.
5. Dennis, W., & Nagarian, P. Infant development under environmental handicap. *Psychol. Monog.*, 1957, **71**, No. 436.
6. Goldfarb, W. Effects of early institutional care on adolescent personality. *Amer. J. Psychiat.*, 1943-4, **102**, 106.
7. ———. Effects of psychological deprivation in infancy and subsequent stimulation. *Amer. J. Psychiat.*, 1945-6, **102**, 18-33.
8. ———. Infant rearing and problem-behaviour. *Amer. J. Orthopsychiat.*, 1943, **13**, 249.
9. Lewis, H. Deprived Children. Oxford: Oxford Univ. Press, 1954.
10. O'Connor, N. Evidence for the permanently disturbing effects of mother child separation. *Acta Psychol.*, 1956, **12**, 174.
11. Pasamanick, B., Rogers, M. E., & Lilienfeld, A. M. Pregnancy experience and the development of behaviour disorder in children. *Amer. J. Psychiat.*, 1956, **122**, 613.
12. ———. Association of maternal and fetal factors with development of mental deficiency. *J. Amer. Med. Assoc.*, 1955, **159**.
13. Spitz, R. A. Hospitalism: An inquiry into the genesis of psychiatric conditions in early childhood. *Psychoanal. Stud. Child*, 1945, **1**, 53.
14. Stott, D. H. Effections of separation from the mother in early life. *Lancet*, 1956, 624-628.
15. ———. Physical and mental handicaps following a disturbed pregnancy. *Lancet*, 1957, 1006-1012.
16. ———. Some psychosomatic aspects of casualty in reproduction. *J. Psychosomat. Rev.*, 1958, **3**, 42.
17. ———. Evidence for pre-natal impairment of temperament in mentally retarded children. *Vita Humana*, 1959.
18. Tredgold, A. F. Mental Deficiency. London: Bailliere, Tindall & Cox (7th ed.), 1949.

Department of Psychology
University of Glasgow
Glasgow, W.2, Scotland

The Journal of Genetic Psychology, 1959, 94, 253-262.

IMMANENT JUSTICE IN CHILDREN: A REVIEW OF THE LITERATURE AND ADDITIONAL DATA*

Institute of Child Welfare, University of Minnesota

GENE R. MEDINNUS[1]

A. INTRODUCTION

In his book, *The Moral Judgment of the Child* (10), Piaget presented a theory concerning the development of moral judgment. In general, he outlined systematic changes with age in children's moral judgments. These changes are a function of age insofar as the child's chronological age defines the nature and quality of his social relationships. Specifically, Piaget emphasized the importance of the progression from the constraint exercised by the unilateral authority of parental discipline in early childhood to a relationship of coöperation existing among peers in the later childhood years.

One aspect of the transition with age in moral judgment is a change in the child's notion of justice. The young child tends to emphasize the rightness and necessity of strict punishment and the importance of submitting to parental dictates; later he is concerned with equality among peers and with types of punishment that more nearly "fit the crime." Piaget examined a number of aspects of the child's developing sense of justice. These include the child's attitude toward the necessity and efficacy of physical punishment; the child's selection of various types of punishment, reflecting his position regarding expiatory punishments as opposed to punishments by reciprocity; his belief in immanent justice; his notions pertaining to distributive or equalitarian justice; and finally, the child's judgment in situations which pit the demands of adult authority against the child's desire for equality of treatment among siblings and peers.

The present paper is concerned only with the child's belief in immanent justice. Piaget attributed the child's belief in the inevitability of punishment to adult constraint. Children are often made to feel that accidental falls and minor injuries are punishments because their parents have said, "It serves you right"; "That will teach you to be careful"; or, "God made it happen to teach you a lesson." Piaget noted elsewhere (8, p. 247) the tendency in children to regard the whole world as aware of their misdeeds,

---

* Received in the Editorial Office on March 12, 1958.

[1] This study was done under the direction of Prof. Mildred C. Templin, Institute of Child Welfare, University of Minnesota.

and their tendency to feel that they are being seen and watched by objects in the Universe, such as the sun and the moon. The child's belief in the moral necessity of the physical laws which govern the Universe is related to his belief that everything in nature is "constructed, intentional, and coherent (9, p. 216)." Since the child also tends to ascribe omniscience to adults, the commands which he receives from them are regarded as a part of the system of laws governing the Universe. Consequently, children are led to believe that nature and objects in the physical environment are accomplices of their parents in guaranteeing the infliction of punishment for any and all wrongdoing. Children's belief in the existence of automatic punishments which emanate from things arises primarily, then, from the unilateral relationship existing between children and adults. Later, through experiences which show that wrongdoing may go unpunished and virtue may go unrewarded, through the discovery of the imperfection of adult justice, combined with the child's intellectual development, belief in immanent justice diminishes.

Piaget (10, p. 251) listed the following percentages of children at the various age levels who affirmed the existence of immanent justice: age 6, 86 per cent; ages 7–8, 73 per cent; ages 9–10, 54 per cent; and ages 11–12, 34 per cent.

B. Review of Related Literature

A number of investigators have attempted to verify Piaget's findings on other social and ethnical groups, and to identify various factors related to the development of relativistic thinking in the moral judgment area. Only the studies and the sections of the studies dealing with immanent justice will be discussed.

In an investigation of moral realism in American children, Lerner (4) obtained value judgments from 389 children in the first six grades of the public school system of an eastern community. Children from two different socio-economic levels of the city were used; one group came from the upper-class residential districts and the other group was drawn from the poorer section populated primarily by people of foreign extraction. Lerner asked one question pertaining to immanent justice. His findings agree with Piaget's in showing a decrease with age in the number of children expressing a belief in immanent justice. This decrease was more pronounced, however, among those children from the better socio-economic group.

Liu (5), at Columbia University, investigated the influence of cultural background on the moral judgment of children. He attempted to define differences in moral realism responses between Chinese- and non-Chinese-American children which could not be accounted for by differences in intelligence and socio-economic status. Liu also found a decrease in belief in immanent justice from age six to 12 but at each age level more non-Chinese children revealed a belief in immanent justice. Liu concluded that decreasing moral realism is not due to age maturity alone.

Abel (1) employed Piaget's technique to study the moral judgments of mentally retarded girls. Comparisons were made between groups in terms of institutionalization, length of institutionalization, and mental age. The subjects were 94 subnormal adolescent white girls with chronological ages ranging from 15 to 21 years and mental ages from 6 through 11. Seventy-four of the subjects were institutionalized at a New York state institution for mental defectives; the other 20 were attending trade adjustment classes in the Manhattan High School for Women's Garment Trades. In the comparison between the institutionalized girls and those living in the community, it was found that approximately twice as many of the former gave responses which indicated a belief in immanent justice. A greater percentage of those institutionalized over six years expressed a belief in immanent justice as compared with those institutionalized under one year. No difference in response was found between two groups equated for chronological age but differing in mental age.

MacRae (6) administered a questionnaire containing moral judgment questions drawn from Piaget's work, and other similar questions, together with a set of questions about the child's social relations (background variables), to a group of 244 boys between the ages of 5 and 14. The questionnaire included two stories designed to elicit the subjects' position with regard to belief in immanent justice. Although in general MacRae found a decrease over the age range in the children's belief in immanent justice, the trend was not consistent. For both stories, a smaller percentage of the five- and six-year-olds gave responses which were scored as indicating a belief in immanent punishment than did the nine- and ten-year-olds.

Although Dennis (2), in his work with Hopi Indian children, was primarily concerned with the concept of animism, he also questioned his subjects regarding their belief in immanent punishment. Ninety-eight subjects, 43 girls and 55 boys, between the ages of 12 and 18 years were examined. Dennis found that 64 per cent of his youngest subjects indicated a belief in the immanence of punishment. This figure decreased markedly with age, standing at 47 per cent for the 14- and 15-year-olds, and dropping to only 9 per cent among the 16- and 17-year-olds. These percentage figures were uniformly higher than those found by Lerner (4) on white children, but were below those obtained by Abel (1) on institutionalized feeble-minded girls of comparable chronological ages. Dennis concluded that the differences in the rate at which early ideas concerning animism are abandoned in Hopi and white communities may be due to a variety of cultural factors which remain, as yet, unidentified. Dennis contended, however, that "the earliest ideas of children are uniform in all societies and are the product of universal child experiences and of mental maturity (2, p. 35)."

As a part of the Indian Education Research Project of the Committee on Human Development of the University of Chicago and the United States Office of Indian Affairs, Havighurst and co-workers (3) studied the moral and emotional development of Indian children. In the overall field project, 902 Indian children from six Indian tribes in the age range from six to 18 years were studied. One part of the larger study is pertinent to the present discussion. Concerning immanent justice, the authors attempted to test Piaget's hypothesis that, in contrast to children in modern societies, children in primitive societies become more rather than less rigid in their moral ideas as they grow older because these societies exercise an increasing amount of moral constraint on them. Thus, if these societies have a view of the world order which includes a supernatural power that watches over men and rewards and punishes their actions, then belief in immanent justice will be as strong in older children as it is in younger children or perhaps even stronger. The findings in this study showed that there was either an increase or no change with age in the children's belief in immanent justice for most of the groups studied. With the exception of one of the Indian groups, the proportion of the 12- to 18-year-olds indicating a belief in immanent justice was approximately 85 per cent. At the younger ages the proportion believing in immanent justice tended to be smaller. These data seem to substantiate Piaget's theory concerning belief in immanent justice in primitive societies. They do not, however, agree with Dennis's (2) findings in his study of Hopi children discussed above in which he found a marked decrease with age in belief in immanent justice.

To test the hypothesis of a relationship between belief in immanent justice and degree of acculturation, Havighurst divided the people from one of the Hopi communities into three groups with differing degrees of acculturation. The three groups were hardly distinguishable in their percentage of immanent justice responses. These investigators concluded that, rather than looking to the concept of primitive cultures subject to varying degrees of influence from a modern culture for an explanation of these results, "we must look for total environmental factors within each group of Indians and of whites to explain the results given by that particular group (3, p. 153)."

In discussing these results, Thompson (11) argued that the basic religious orientation of these people concerning the nature of the Universe and man's relation to it is deeply rooted in the tribal past and persists regardless of changes in the group's economy or exposure to white cultures. This would seem to account, at least in part, for the lack of relationship between the degree of acculturation and belief in immanent justice.

In sum, a number of investigators have sought to verify and extend Piaget's findings concerning changes with age in children's belief in immanent justice. These studies have supported Piaget's data regarding a decrease in belief in immanent justice with age. In a comparison of children from two different socio-economic levels, Lerner found this decrease to be more pronounced among the children from the better socio-economic group. Liu, contrasting Chinese- and non-Chinese-American children, reported that more non-Chinese than Chinese children revealed a belief in immanent justice at each age level from six to 12. Abel's comparison of the responses of two groups of mentally-deficient girls showed little relationship between mental age and belief in immanent justice. Abel interpreted the differences which she found between institutionalized girls and girls living in the community as indicating the constraining influence of an institutional atmosphere. Though MacRae found a decrease in belief in immanent justice from age seven to 14, this trend was reversed at the five- and six-year levels where the percentage of subjects expressing this belief was comparable to the 11- to 12-year level. Dennis found a higher proportion of Hopi children expressing a belief in immanent punishment than is usually found among white children. The decrease with age in this belief which Dennis found was not supported by Havighurst's findings in a study of six American Indian tribes.

C. PROCEDURE

1. *Description of the Subjects*

Two hundred and forty children, 30 boys and 30 girls, at each of the following four age levels, 6, 8, 10, and 12, served as subjects in the present study. The subjects were selected from the lower socio-economic groups, specifically from Classes V, VI, and VII according to the Minnesota Scale for Paternal Occupations (7). Children from these groups were used so that the sample would be as comparable as possible to that used by Piaget who studied children from "the poorer parts of Geneva" (10, footnote, p. 37).

2. *Method*

Eighteen story-situations and questions drawn from Piaget were administered to each subject individually. The subjects' responses to the questions were recorded verbatim by the examiner. Only those two items pertaining to immanent justice will be discussed in this paper. The two stories which were employed to elicit the subjects' position with regard to belief in immanent justice were:

Story I

There was once a little boy who didn't mind his mother. He took the scissors one day when he had been told not to. While he was trying to cut some paper, he cut his finger.

Why did he cut his finger?

If his mother had said it was all right if he used the scissors, would he have cut himself just the same?

Story II

Once there were two children who were walking by a house in the country. There were some apple trees out in the yard in front of the house. No one was around so they went into the yard and stole some apples. Suddenly a man came out of the house and ran after them. He caught one of the boys but the other one got away. This one crossed a river on a rotten bridge and he fell into the water.

Why do you think the boy fell into the water?

If he had not stolen the apples but he had crossed the river on that rotten bridge, would he have fallen into the water anyway? Why?

Following Lerner's method of scoring, a plus was given those responses which indicated that the subject conceived of the accidents described in each of the two stories as forms of punishment administered by the objects themselves, or by God, or by the "moral force" in the Universe through these objects. A minus score was assigned those responses which involved a rational explanation; for example, a minus was given those responses to Story II which stated that the boy fell into the river because the bridge was rotten.

D. RESULTS

The subjects' responses to the two stories are tabulated in Table 1. Although the number of subjects revealing a belief in immanent justice decreased with age for Story I, the opposite trend is indicated for Story II. The chi-squares for the two stories are not significant, however. The sex difference for the stories was likewise insignificant.

TABLE 1
NUMBER OF + AND − VALUATIONS AT EACH AGE LEVEL FOR STORIES I AND II

Age	Story I		Story II	
	+	−	+	−
6	37	23	16	44
8	32	28	15	45
10	30	30	23	37
12	27	33	23	37
	$\chi = 3.57$ ns		$\chi = 4.36$ ns	

E. DISCUSSION OF RESULTS

Whereas Piaget found that 86 per cent of the six-year-olds whom he questioned affirmed a belief in immanent justice, only 62 per cent of the children at this age level in the present study indicated such a belief in response to Story I. An examination of the answers which the children gave in response to the question, "Why did he cut his finger?" provides a partial clue to the comparatively small percentage found in this study. The rational explanations offered by the six-year-olds seemed to indicate that they were drawn from the child's personal experiences. The following are representative of this type of response: "Because he might not was careful." "Maybe he was too little." "Maybe he got his finger too close to the scissors." "Maybe they slipped." "The scissors were sharp." "Because he didn't watch what he was doing." The situation described in the story seems to be one which is well within the experience of the children at this age level. In other words, the children did not need to invoke the operation of a supernatural force to account for the misfortune which befell the boy in the story. The mishap could be accounted for on the basis of a commonplace experience.

An examination of Table 1 shows that a smaller percentage of children at each age level evinced a belief in immanent justice for Story II than did so for Story I. A comparison of the contents of the two stories throws some light on this discrepancy. In Story I no rational explanation is provided to account for the boy's misfortune which occurred while he was in the act of cutting up some paper. In Story II, on the other hand, mention is made of the fact that the bridge on which the culprit crossed the river was rotten. In the latter story, then, the subject need not resort to an explanation of the punishment inflicted based on the intervention of a supernatural force operating through objects in the environment; rather, he can utilize the information presented to him to account for the incident. It would seem that the tendency for children to explain occurrences on the basis of a supernatural vigilance and conspiracy depends in part upon the amount of information which is available to them in a given situation.

The results of the present study do not corroborate Piaget's conclusion that there is a marked decrease with age in the number of children who believe in immanent justice. Although it is apparent that the percentage figures differ from one story to another, Piaget did not tabulate separately the result obtained from different stories, and, in fact, it is not clear from which stories the results given were obtained. A similar criticism might be leveled at the other studies (1, 2, 3, 4) which employed only one story in an attempt to determine the number of children at various age levels who believe in immanent justice. The present findings indicate that the point of view which a child expresses varies depending on the story used so that any percentage figure is specific to a particular story. Before one could speculate or theorize about differences in the belief in immanent justice among various cultural, ethnic, and other such groups, one would have to be concerned about the differences obtained from the same group by stories which vary in content. An attempt to specify precisely the reasons for this would seem to be the most fruitful problem for future research in this area.

In summary, the present findings indicate that a child's expressed belief in immanent justice is dependent upon a number of such factors as the meaningfulness of the situation to him, the presence or absence of rational alternative explanations, the range of his experiences, the concreteness of the young child's thought, and so forth. The possible importance of specific parental teachings will be discussed briefly in the next section.

F. Suggestions for Further Research

It was found that a number of six-year-olds responded to Story I (a child cuts himself with scissors) on a very concrete level; the incident described was within their range of experience. It might be postulated that whether or not a child demonstrates a belief in immanent justice depends upon whether or not the situation described is within his range of experience and understanding. It would follow that the more foreign and remote the situation is from the child's experience, the more likely it will be that he will give an immanent justice type of response. To test this, two groups of items could be developed: in one group, the incidents described would be placed in contexts meaningful to the child; the incidents in the other group would be placed in contexts alien to the child's experience. The thesis is that children tend to explain events in supernatural terms when the situation is unfamiliar to them.

The present study also found that children tend to be less likely to give supernatural explanations of incidents in stories where an alternative rational explanation is provided. It was argued that whether or not a child invokes the operation of a supernatural force to explain an event depends upon the amount of information which is available to him. To test this, stories which are essentially similar except for the presence or absence of a rational explanation could be administered to a group of children, and the number of children expressing a belief in immanent justice for the two types of stories could be compared.

Although Piaget seemed unclear in his position regarding the importance of adult teachings in producing an "immanent justice" type of thinking in children, it would seem that parental attitudes in this respect are significant determiners of the child's tendency to explain punishment and other such events in this manner. A possible method to get at such a relationship would be to formulate subtle items appropriate to the adult level in order to assess the parents' way of thinking. It would be hypothesized that adults do differ in this regard and that this influences the child's tendency to consider punishments as automatic and inevitable.

G. Summary

A review of the literature concerning Piaget's notion of immanent justice revealed that the various studies supported Piaget's findings which showed a decrease with age among children in belief in immanent justice. Several environmental factors affecting this belief were identified.

In the present study, 240 children, 30 boys and 30 girls, at each of the following four age levels, 6, 8, 10, and 12, were administered a questionnaire consisting of stories drawn from Piaget's work. The subjects were selected from the lower socio-economic levels. The data obtained from the two immanent justice items included in the questionnaire were presented and discussed.

Although the number of subjects revealing a belief in immanent justice decreased from age six through age 12 for Story I, the opposite trends was found for Story II. Tentative explanations for these results were found in the nature of the stories themselves. It was concluded that a child's expressed belief in immanent justice is dependent on such factors as the meaningfulness of the story described, the presence or absence of rational alternative explanations, and the range of the child's experiences.

Several suggestions for further research in this area were discussed.

REFERENCES

1. ABEL, T. M. Moral judgments among subnormals. *J. Abn. & Soc. Psychol.,* 1941, **36**, 378-392.
2. DENNIS, W. Animism and related tendencies in Hopi children. *J. Abn. & Soc. Psychol.,* 1943, **38**, 21-37.
3. HAVIGHURST, R. J., & NEWGARTEN, B. L. American Indian and White Children, a Sociopsychological Investigation. Chicago: Univ. Chicago Press, 1955.
4. LERNER, E. Constraint Areas and the Moral Judgment of Children. Menasha, Wisc.: Banta, 1937.
5. LIU, CHING-HO. The influence of cultural background on the moral judgment of children. Ph.D. dissertation, Columbia Univ., 1950.
6. MACRAE, D., JR. The development of moral judgment in children. Ph.D. dissertation, Harvard Univ., 1950.
7. MINNESOTA SCALE FOR PATERNAL OCCUPATIONS. (Publ. of Inst. of Child Welfare.) Minneapolis: Univ. Minnesota Press, 1950.
8. PIAGET, J. The Child's Conception of the World. New York: Harcourt, Brace, 1928.
9. ———. The Language and Thought of the Child. (2nd ed., trans. M. Gabain.) New York: Harcourt, Brace, 1932.
10. ———. The Moral Judgment of the Child. (Trans. M. Gabain.) Glencoe, Ill.: Free Press, 1948.
11. THOMPSON, L. Attitudes and acculturation. *Amer. Anthropol.,* 1948, **50**, 200-215.

229 East Cherry Street
Watseka, Illinois

The Journal of Genetic Psychology, 1959, **94**, 263-266.

THE INFLUENCE OF AUDITORY STIMULATION ON ANISEIKONIC PERCEPTION: II AND III. BUSINESS SCHOOL WOMEN AND COLLEGE WOMEN*

The Menninger Foundation

JAMES E. SIMPSON[1, 2]

The series of studies of perceptual distortion with aniseikonic spectacles completed in the Menninger laboratories consisted of three related but independent experiments. All three experiments were similar: designed to test autism hypotheses (3), based on the same operational rationale, and carried out with the same equipment and experimenter and similar subjects and procedures. The basic plan of these experiments is described in detail in the report of the preliminary experiment (4) and is summarized here.

The experimental plan in all three studies was to investigate the influences of induced-affects (via auditory stimulation) on aniseikonic effects of (*a*) distortion of the experimental room and (*b*) reaction time (seconds for distortion to occur). It was postulated that aniseikonic effects associated with the pleasant music would be accepted more readily (*decreased* reaction time and *increased* distortion) than aniseikonic effects associated with the unpleasant noise (*vice versa*). Young women were individually given a pseudo-test of proficiency in adjusting a selsyn-controlled measuring-arm parallel to the backwall of the room while wearing aniseikonic spectacles. Changes in reaction time and degrees of distortion in association with pleasant music and upleasant noise were obtained.

The complex results of the preliminary experiment (4) cannot be clearly analyzed because the confounded factorial design does not allow for separation of the counterbalanced paired-variables (two sets of lenses vs. two stimulus-orders). Nonetheless, the reaction-time data show that generally the aniseikonic distortion previously associated with pleasant music was more quickly perceived than the distortion previously associated with unpleasant noise (17 *S*s positive, 7 *S*s negative, and 4 *S*s neutral; chi-square = 4.17,

---

* Recommended for publication by Dr. Gardner Murphy, and received in the Editorial Office on March 18, 1958.

[1] In addition to the support of The Menninger Foundation, this investigation was supported by a research grant (M-715C) from the National Institute of Mental Health of the National Institutes of Health, Public Health Service.

[2] The author is grateful to Drs. Gardner Murphy, Riley W. Gardner, and Wayne A. Holtzman, for continued encouragement and guidance.

$p < .05$). The distortion data are not significant (17 *S*s positive and 11 *S*s negative).

The second experiment was designed to eliminate confounding of factors and allow for analysis of covariance. Simplicity in design was obtained by using only one set of spectacles (1, p. 337, Figure 1, *D*) for all *S*s and one auditory stimulus, either pleasant or unpleasant, for each experimental *S* (thus, the confounding of two sets of lenses vs. two stimuli vs. two orders was avoided). Each of the 39 *S*s was given three sets of five trials: pre-training, training, testing. The pre-training trials were the same for all three groups (music-training vs. noise-training vs. control group). Individual differences in the pre-training trials were held constant via analysis of covariance. The group results and the individual results for the reaction time data for the 39 *S*s are not as predicted (21 *S*s showed zero-seconds reaction time for most of the trials). The overall analyses indicate that generally the music-trained *S*s increased in degrees of distortion and the noise-trained *S*s decreased, as predicted (*t* test of adjusted means, $t = 2.48$; $df = 23$; $p < .01$). However, the adjusted mean of the music-trained group is not significantly different from that of the control group, suggesting the need for replication with larger samples. The individual results show that the music-trained group increased in distortion, the control group remained relatively unchanged, and the noise-trained group decreased in distortion, as predicted (Table 1).

TABLE 1

Stimulus group	No. *S*s increasing in degrees distortion	No. *S*s decreasing in degrees distortion	Mean % change in degrees distortion
Music-trained	9	4	16% Increase
Control	6	7	4% Increase
Noise-trained	4 (1 *S* no change)	8	13% Decrease

The chi-square between increasing distortion and decreasing distortion was significant at less than the .05 point.

The third experiment replicated the second experiment in that the first 15 of the 25 trials were the same: pre-training, training, testing. The additional 10 trials were another series of training and test trials and (like the first experiment) provided both music-training and noise-training for each *S*. Also, after the 25 trials each of the 72 *S*s rated the affective potency (pleasantness vs. unpleasantness, distraction, and attention values) of the music and the noise on rating scales and thereby provided data for assessment of the intervening variable, experimentally-induced affect.

The reaction time data of the third experiment were similar to those of the second (49 of the 72 *S*s had zero-seconds reaction time for most of the trials) and are not significantly related to any specific hypothesis. Support of the basic rationale of the experiment is the significant correlation ($r = - .26$, $p < .02$) between log of initial reaction time(in *S*'s first trial with aniseikonic spectacles) and mean degrees distortion for the control trials (the first five trials with aniseikonic spectacles); i.e., the quicker the *S* sees the distortion effect the greater amount of distortion seen. This compares with r of $- .39$ and $p < .01$ for these two scores in the second experiment. The correlation for the combined group of 111 *S*s in the second and third experiments is $- .32$, $p < .001$.

The distortion data of the third experiment showed no significant relationship to any predictions except as noted in the preceding paragraph. In fact, the individual scores for mean degrees distortion are almost perfectly balanced for and against the hypotheses; for example, Music changes: 42 *S*s for hypotheses and 30 against vs. Noise changes: 30 for, 41 against, and one *S* no change in distortion. The regression scores—residual error—show no such differential responses of *S*s to music-training and noise-training: 64 of the 72 *S*s either increased distortion after both music-training and noise-training or decreased distortion after both. Further, these residual scores for distortion changes in response to music and noise are highly correlated ($r = .79$). The consistent, marked individual differences are not due to order and/or differential-stimulus effects (in a latin-square analysis of music vs. noise distortion, the only significant F test was that between *S*s: $p < .001$). The aniseikonic effects data of the second and third experiments are presented in Table 2.

Another indication of the consistency of the individual *S*s throughout the 25 trials of Experiment III is the modified odd-even (second plus fourth

TABLE 2

REACTION-TIME AND DEGREES-DISTORTION DATA FOR 111 YOUNG WOMEN IN CONTROL TRIALS OF EXPERIMENTS II AND III

	Initial Reaction time*				Mean degrees distortion**			
	Lowest	$\overline{X}$	Sigma	Highest	Lowest	$\overline{X}$	Sigma	Highest
Experiment II (39 business school students)	9.0	104.8	74.22	442.0	7.1°	19.03°	7.38°	41.0°
Experiment III (72 sorority coeds)	7.0	101.5	65.38	614.0	4.2°	18.22°	6.59°	37.2°
Total (111 *S*s)	7.0	102.24	69.29	614.0	4.2°	18.51°	6.89°	41.0°

* Seconds required by *S* to experience distortion in first trial.

** Mean degrees distortion for the five trials.

series of trials, i.e., music and noise trials, vs. third plus fifth series of trials, i.e., silent trials) reliability coefficient of .94 for "adjustment time" (seconds required by *S* to adjust selsyn-controlled measuring-arm parallel to the backwall). The consistency of the individual *S*s in Experiment III is especially important because the same 72 *S*s were variously tested in other perception experiments in the Menninger laboratories, and the comparative data so far studied indicate some interesting inter-relationships (2).

The rating-scale data of the third experiment show no significant relationships of the *S*s' affective ratings of the music and noise and the aniseikonic effects of reaction time and distortion. Although almost all of the *S*s spontaneously described (and later rated) the music as pleasant and the noise as unpleasant, their experimentally-induced affects—as reflected in their ratings—are not related to their aniseikonic responses.

The predicted results of differential changes in distortion by the music-trained *S*s vs. the noise-trained *S*s in the second experiment were not obtained by replication in the third experiment. This lack of cross-validation may be accountable to social differences in the two samples. The *S*s of the second experiment (as in the first) were younger women enrolled in business school and were mostly recent high-school graduates from small rural communities. The *S*s of the third experiment were all sorority women enrolled as freshman-to-senior university students and were mostly from the urban area. The sorority women were older (19.32 years vs. 18.93; $t = 2.05$; $df = 96$; $p < .05$ level) than the secretarial students, but age per se seems less important than the more obvious differences in sophistication. The *E* tested individually each of these 111 *S*s; the naïveté of the secretarial students was in marked contrast to the poised alertness of the sorority women. It may be that the more impressionable secretarial students could be experimentally manipulated but the more sophisticated sorority women were resistive to strong experimentally-induced affect via music and noise.

REFERENCES

1. AMES, A., JR. Binocular vision as affected by relations between uniocular stimulus-patterns in commonplace environments. *Amer. J. Psychol.*, 1946, **59**, 333-357.
2. GARDNER, R. W., SIMPSON, J. E., & SNYDER, C. W. (MS in preparation.)
3. MURPHY, G. Affect and perceptual learning. *Psychol. Rev.*, 1956, **63**, 1-15.
4. SIMPSON, J. E. The influence of auditory stimulation on aniseikonic perception: I. A preliminary study. *J. of Psychol.*, 1956, **41**, 235-241.

Department of Psychology
Oregon State College
Corvallis, Oregon

The Journal of Genetic Psychology, 1959, **94**, 267-285.

RORSCHACH INDICES OF DEVELOPMENTAL LEVEL*[1]

Worcester State Hospital and Clark University

LESLIE PHILLIPS, STANLEY KADEN, AND MARVIN WALDMAN[2]

A. INTRODUCTION

For the past seven years we have been exploring the interrelationships of perceptual function and both socially adaptive and deviant behavior patterns. In the course of this work we have focused upon two major areas. The first has been the development of a systematic method for the comparative study of normal and deviant perceptual activity. This method has been validated and applied in a series of studies on normal and psychopathological groups. This paper will briefly review these investigations and present a revision of the method for the assessment of perceptual function. Succeeding papers will deal with our second area of interest, i.e., an analysis of the cognitive basis of those aspects of perceptual function measured by our revised perceptual indices.

We shall first report on a series of validation studies on an earlier form of this system. This will be followed by a survey of some investigations of normal and psychopathological behaviors utilizing this method. Finally, we shall present the current revision of the method for assessment of perceptual activity. All these studies have been undertaken within the framework of general developmental theory as expounded by Werner (34) and Piaget (26). Werner has stated that "developmental psychology seeks to coördinate within a single framework, *forms of behavior* observed in comparative animal psychology, in child psychology, in psychopathology, in ethnopsychology, in general and differential psychology of men of our own culture" (35). The general principle evoked is that "wherever development occurs it proceeds from a state of relative globality and lack of differentiation to a state of increasing differentiation, articulation, and hierarchic integration. This principle has the status of an heuristic definition. Though itself not subject to empirical test, it is valuable to developmental psychologists in leading to

---

* Received in the Editorial Office on March 24, 1958.

[1] This investigation was supported by the Dementia Praecox Research Project, Worcester State Hospital, and a research grant (M-896) from the National Institute of Mental Health, Public Health Service.

[2] Thanks are extended to the members of the Psychology Department, Worcester State Hospital for their many constructive criticisms in the writing of this manuscript.

a determination of the actual range of applicability of developmental concepts to the behavior of organisms" (35).

Many investigators have attempted to utilize a general developmental model in their approach to the understanding of human behavior (11, 22, 32). Until recently, however, there was little in the way of specific techniques for the comparative study of widely diverse populations. Thus, individuals and groups could be hierarchized developmentally only in terms of such indirect indices as chronological age (of children), normal versus abnormal states, and primitive versus civilized societies. The creation of broadly applicable measures for the direct assessment of psychological and social maturity permits a formal comparison of function in widely diverse populations. The existence of such indices is of significance in the study of psychopathology, normal patterns of social adaptation, ego psychology, and the like.

The Rorschach technique is a particularly appropriate device for the assessment of perceptual behavior. The ambiguity of the stimulus conditions allows freedom for great diversity in responses and yet the nature of the task is such that these reactions may fairly easily be ordered into formal categories on a wide range of characteristics. In addition, the applicability of the test to a variety of populations makes it a most suitable technique for studies utilizing the developmental approach.

Dworetzki (3) was the first to utilize the Rorschach technique in this fashion. She devised a schema for the genetic ordering of Rorschach test responses and applied her measures of perceptual activity in studies of normal children ranging in age from 2 to 15 years, retarded children, educated and uneducated adults, and schizophrenic adults, and reported on the responses of these groups. She found that genetically low subjects, including young and retarded children, uneducated adults, and schizophrenics shared many characteristics indicative of perceptual inadequacy. Older children tended increasingly to perform as do adults.

Friedman (9) modified and quantified the scoring schema of Dworetzki, drawing upon the work of Rapaport and Beck, both of whom have emphasized the variety of psychological processes subsumed under various Rorschach scoring categories. His measures were restricted to location scores, which he proposed as the best representatives of the formal or structural aspects of perceptual activity. The measures were framed explicitly within the basic principle of Werner that development proceeds in terms of increasing differentiation and hierarchic integration. Genetically low whole and detail responses "are those which indicate amorphous, diffuse, or confabulatory percepts where little attention is given to part relations and to perception of contours. The corresponding genetically high percepts are reflected in the responses whereby the percept is that of a precisely formed unit with parts integrated into it and where the whole is composed of independently perceived details organized into an integrated totality" (35).

The reliability of this scoring schema has been reported extensively in various studies (9, 13, 18, 31). In these studies the percentage agreement among independent scores ranged from 89.7 per cent to 95.6 per cent, with an approximate mean as reported by Lane (18) of 93 per cent.

Validation procedures were dictated by the nature of theory. The abstract formal emphasis of the developmental approach calls for validation not only in terms of chronological age in children, but also in terms of other aspects of behavior which may be developmentally ordered. Thus, three types of validation studies were undertaken in order to test the proposition that scoring on these Rorschach indices reflects developmental level of functioning. These studies involved an examination of the relationships of these measures to chronological age in children (13), increasing tachistoscopic time of exposure of the Rorschach cards (6), and degree of psychiatric impairment (9, 23, 29, 31).

Hemmendinger (13), in a study of 169 children between the ages of 3 and 11 years, confirmed the hypothesis that higher scoring, developmentally considered, on these Rorschach indices parallels increasing chronological age. In the tradition of Sander's (30) "microgenetic" work on the emergence of percepts, Framo (6) administered the Rorschach test tachistoscopically at .01 of a second, .10 of a second, 1.0 second, and 10.0 seconds to 80 normal adults, 20 subjects at each exposure time. He found, using the same Rorschach scoring system as Hemmendinger, that genetically high types of whole and usual detail responses became more frequent with increased opportunity to utilize time, genetically early responses tended to drop out sharply as exposure time is increased, and the general tendency is in the direction of maturity of percept and integration of blot areas. A comparison of the responses in this study by Framo with the ontogenetic data obtained by Hemmendinger show striking agreements.

A series of validating studies based on perceptual functioning in psychiatrically impaired groups was also carried out (9, 23, 29, 31). The hypothesis was that degree of psychiatric impairment would parallel genetic level of Rorschach responses such that the most severely impaired patients would show the genetically earliest responses. The results of these studies may be summarized as follows: hebephrenic and catatonic schizophrenics manifested perceptual activity formally similar to children from 3 to 5 years of age while paranoid schizophrenics paralleled children from 6 to 10 years of age. A group of psychoneurotics placed "intermediate and overlapping" with the performance of 10-year-old children on one hand and normal adults on the other. The perceptual performance of cerebrally-damaged adults resembled that of a group of young children more closely than that of a group of older children. Mentally defective children were found to manifest more primitive perceptual processes than either normal children or adults of varying degrees of psychiatric impairment.

From the studies reported above it can be concluded that genetic level of perceptual behavior is susceptible to quantification. Further, there is support for the assumption that developmental level in this area of psychological function is related to significant personality and behavioral parameters.

The utilization of explicit measures of maturational level within the context of a general developmental theory which may serve to organize formal characteristics of normal and pathological behavior patterns has led to a further series of investigations. One of these studies has sought to relate developmental formulations to the adaptive adequacy of normal adult men. Lane (18) has recently investigated the relation of developmental level of perceptual function to level of social maturation, as measured by the Worcester Scale of Social Attainment (24). Lane noted that an increasing emphasis on planful action and on emergence of a sense of personal responsibility can be taken to parallel cognitive development. On the basis of such parallels, Lane hypothesized a relationship between developmental level of cognitive function, measured in terms of the revised perceptual indices and social maturation. His results supported the hypothesis.

In a similar study, Fowler (5) undertook a developmental analysis of perceptual function in normal and schizophrenic subjects. From each of these two populations he chose groups of *S*s matched for low social adequacy, and two similar groups were matched in terms of moderate social adequacy. He found that scoring on Rorschach measures associated with a low level of perceptual maturation discriminated along the dimension of social adequacy, regardless of psychiatric state. In contrast, perceptual measures indicative of the degree to which a high developmental level of perceptual function had been attained differentiated between these normal and psychotic groups, independent of level of achieved social adequacy.

In another study, Lofchie (19) analyzed performance under distraction stress in terms of developmental theory. He drew a parellel between the factors necessary for the production of perceptually mature responses on the Rorschach test and for adequate performance under conditions of distraction.

The common process, he felt, was an integrative cognitive function which allowed the individual to attend to relevant stimuli and at the same time to exclude those stimuli which were irrelevant to the task at hand. He found that ability to withstand distraction was positively associated with developmental level of perceptual activity on the Rorschach test. Similar findings were obtained by Waldman (33).

In a recent study of creative individuals, Hersch (14) found support for the hypothesis that highly creative artists, as compared with more typical normal individuals, are characterized by the greater availability of relatively primitive cognitive functions and particularly by a greater emphasis on more mature cognitive operations, as measured by the Rorschach developmental indices.

Several recent studies have been concerned with some of the major dimensions along which psychopathological phenomena may be ordered. In addition to the group of investigations which related severity of psychiatric impairment to level of perceptual function reviewed above, a study by Freed (8) involved a comparison of the perceptual process in the severely psychotic and in normal subjects by means of tachistoscopic procedures. Freed hypothesized that, in contrast to normal subjects, hebephrenic and catatonic schizophrenics would fail to show increased perceptual differentiation with increased exposure time. Using the same technique as Framo, he showed the Rorschach cards at .01″, .10″, 1″, and 10″ to a group of 60 hebephrenic-catatonic schizophrenics, 15 at each exposure time. Structurally, their performance was not grossly different from that of normal adults at the shortest time of exposure, but as this was lengthened the schizophrenics lagged increasingly in the development of perceptually mature responses.

Another series of papers has been concerned with the dimension of displacement and thus with vicarious modes of symptom expression. In these studies it was theorized that motoric symptoms are manifestations of a genetically lower psychological organization than is ideational symptomatology. As a test of this proposition, Hurwitz (15) studied two groups of children, the major symptomatology in one group being underactivity, in the other hyperactivity. As predicted, individuals in the underactive category scored genetically higher than the acting-out group. Misch (20) compared individuals who had a history of only threatening assault with a group who were classed as chronically assaultive. Kruger (17) likewise compared individuals who threatened suicide with those who made an actual attempt. She also compared overtly sexually-deviant individuals with a group who were only concerned over the possibility of such aberrent sexual behavior. In all these studies individuals who tend to "act out" their symptomatology rather than to express it ideationally perform at a developmentally lower perceptual level, i.e., manifest a less mature pattern of cognitive organization.

The hypothesis that regression of cognitive function occurs in adolescence and that this will be reflected in perceptual organization was tested and confirmed by Wilson (36). A similar proposition in regard to cognitive primitivation in aged men was supported by the findings of Rochwarg (28).

While the results of all the studies affirmed the efficacy of Rorschach indices as appropriate measures for the assessment of developmental level of perceptual functioning, an extension and revision of this scoring system was deemed necessary for several reasons. First, the original form of the system dealt only with location scores. Recently we have found that other aspects of Rorschach performance (e.g., movement and degree of form dominance) also serve to order developmentally various laboratory behaviors as well as normal and pathological reactions. Further, in the original form of the system, some scores failed to differentiate between groups known to differ in developmental level by some external criterion. Consequently, these measures have been further refined and other scores have been combined into single indices. These extensions and revisions were made on the basis of empirical data. For example, the occurrence of movement response and the degree of form dominance have been reported by a number of investigators (4, 16) to show a correspondence with increasing chronological age in children. Further, certain Rorschach records provided by Witkin[3] in a study of field dependence served as the basis for the development of the functional integrative measure reported below. Finally, in a number of our studies, the Rorschach responses which serve as the basis of the developmental scoring schema have been scrutinized to see in what manner response classification might be modified to sharpen obtained relationships. Such changes in the scoring system have been introduced in subsequent studies. It seems advisable to publish this newly revised scoring system, particularly since independent investigators (2)[4] have recently reported on research which makes use of earlier systems of response classification.

---

[3] Personal Communication.

[4] Also: Wilensky, H. The relationship of Rorschach developmental level to hospital adjustment and social participation among chronic schizophrenics. Paper read at Eastern Psychol. Assoc., New York: April, 1957.

B. The Revised Rorschach Developmental Scoring System

The revised developmental scoring system contains rules for scoring four aspects of Rorschach performance.[5]

Adequacy and specificity of form responses: "mediocre," "minus," "vague," and "amorphous."

Determinants used in combination: form dominance versus form subordination.

The perception of activity: the movement responses: M, FM, and m.

Organization of blot elements:

- Adequate organization: the "integrative" responses.
- Inadequate organization: "fabulized combination," "contamination," and "confabulation."

1. *General Rules*[6]

1. Determinants which are developed only in the inquiry but modify the original percept are scored. An example is:

Free Assoc.	*Inquiry*	
1. W Bat	A black bat and it's flying	Wm C'F·FM+ A P

2. A response is scored twice if:

(*a*) There are at least 2 scorable contents which occur in the free association, *and*

(*b*) The 2 scorable contents are not identical

VIII. 2 animals (D1)[7] climbing over some rocks   Wi FM+ A P
Dv F+ Ls

(*c*) If one of the contents is minus form level

VIII. 2 caterpillars (D1) crawling over some shrubbery   W— FM— A
Dv F+ Ls

A response is *not* scored twice if:

(*a*) A new percept which does not modify the original percept is developed in the inquiry.

Free Assoc.	*Inquiry*	
VIII. 2 animals (D1)	They're climbing over some rocks	Wi FM+ A P

(*b*) 2 or more contents occur in the free association but one of the contents occupies approximately less than ¼ of the area of the figure percept, contains only form as a determinant, and falls into the class of "mediocre" response.

III. 2 men (D1) stirring a pot (D7)   Di M+ H P

---

[5] A fifth area of response classification, that of content, is currently being analyzed by Grace (12). The results of that investigation will be reported in a subsequent paper.

[6] For any more general or basic scoring rules than are found in this outline, reference should be made to Phillips and Smith (25).

[7] The numbering of details follows that given by Beck (1).

2. *Adequacy and Specificity of Form*

"Mediocre" is scored when *all* four of the following criteria are met:

(*a*) there is a scorable single content including those responses given in a plural form, e.g., "animals." (For restrictions on this definition see also "integrative" responses.)

(*b*) the content has a specific structural requirement. (See "vague" responses.)

(*c*) the response is of adequate form level (F+).

(*d*) the response does not cross color (including white space) boundaries. (See "Structural Integrative" response.)

The majority of all responses will probably be scored in the "mediocre" category, for this includes all of the easily achieved, common W, D and Dd responses.

Some examples of the "mediocre" response are:

Free Assoc.		*Inquiry*			
I. W	Bat	It's black	Wm	FC′+	A P
III. D3	Red Bowtie	The color and shape	Dm	CF+	Pr P

"Minus" is scored when both of the following criteria are met:

(*a*) there is a scorable single content including those responses given in plural form, e.g., "animals." (For restrictions on this definition see also "Fabulized Combination" and "Contamination".)

(*b*) the response is of inadequate form level (F—).

Some examples of "minus" responses are:

Free Assoc.		*Inquiry*			
II. W	Insect		W—	F—	A
VIII. D1	Caterpillar	Shape and multicolored	D—	FC—	A
IX. Dd21	Teeth		Dd—	F—	Hd

The criteria for inclusion in the class of "vague" responses are as follows:

(*a*) the content must be F+, and

(*b*) the content must not imply a specific structural requirement, e.g., "harbor" implies a *U*-shaped structure and so is *not* scored "vague," whereas "water" has no such specific structure requirements and is therefore scored as "vague."[8]

Some examples of the "vague" response are:

I. W	Clouds	Wv	F+	Cl
II. D3	Explosion	Dv	F+	Expl.

---

[8] Phillips and Smith (25) have ordered all contents, by class, along a continuum of specificity. By making reference to these classes, one can judge any response according to the specificity criterion for inclusion in the "vague" category. Almost all contents in the first three classes of the Phillips-Smith system are scored "vague."

In addition, such contents as shrubbery, stain, map (unelaborated), land, smoke (non-columnar), and rocks are scored "vague" as well as primitive organisms such as amoeba. Anatomy contents cover the entire range of the specificity continuum, and thus must be judged individually. The following anatomy contents, however, if unelaborated, are scored "vague": Medical chart, inside of body, X-ray.

"Amorphous" is scored when *both* of the following criteria are met:

(*a*) the percept falls into the class of "vague" responses

(*b*) any of the following is given as the only determinant: C, C′, c, V, m
Examples are:

Free Assoc.			*Inquiry*				
II.	D3	Blood	The color red	Da	C	Bl	
IV.	W	Something disintegrating		Wa	m	Ab	

3. *Determinants used in Combination*

Usually, the basis for scoring "form dominance" in color, achromatic color, or texture responses is dependent on the inquiry.

The use of color or achromatic color in the free association or the inquiry to *locate* percepts does not influence scoring. Examples of responses where no color is scored: The red part—could be a butterfly, The green looks like an insect, The darker grey part looks like clouds.

Form is scored as dominant if in the inquiry the *S*:

(*a*) mentions color, achromatic color, or texture spontaneously subsequent to details, giving emphasis to shape. Examples are:

Free Assoc.			*Inquiry*				
III.	D3	Butterfly	wings and body, it's red	Dm	FC+	A	
I.	W	Bat	shape, it's black	Wm	FC′+	A	P
VII.	W	Clouds	general formation, it's fluffy	Wv	Fc+	Cl	

(*b*) mentions color, achromatic color or texture only on question. Examples are:

Free Assoc.			*Inquiry*				
III.	D3	Butterfly	The shape (?)—it's red—	Dm	FC+	A	
I.	W	Bat	wings and body (?)—it's black—	Wm	FC′+	A	P
VII.	W	Clouds	shape or clouds (?)—also fluffy	Wv	Fc+	Cl	

Form is scored as subordinate if in the inquiry:

(*a*) the color, achromatic color or texture are given primacy. Examples are:

Free Assoc.			*Inquiry*				
III.	D3	Bow tie	red bow tie	Dm	CF+	Pr	P
VI.	D1	Bear skin	It's furry like a skin	Dm	cF+	Ad	P

(*b*) color, achromatic color or texture mentioned first. An example is:

Free Assoc.		*Inquiry*			
IX. W	Flower	the colors and the shape	Wi	CF+	Bt

(*c*) if color, achromatic color or texture only is used to define a percept, CF, C'F, or cF is scored if that percept falls in the class of "mediocre" responses. Examples are:

Free Assoc.		*Inquiry*			
IX. D1	Tree	a tree because it's green	D—	CF—	Bt
I. W	Bat	because it's black like a bat	Wm	C'F+	A P

If no form is verbalized and the content falls into the class of vague responses, the rules for scoring "amorphous" responses are followed.

Note: If color, achromatic color or texture as a modifying word is used in the free association, *and* none of these determinants is elicited in the inquiry:

(*a*) scoring is FC, FC', or Fc if the percept falls into the class of "mediocre" responses. Examples are: FC: "red butterfly," "green insect," FC': "black bat," Fc: "furry animal."

(*b*) scoring is CF, C'F, or cF if percept falls into the class of "vague" responses. Examples are: CF: "colored painting," "pink clouds," C'F: "black thunderclouds," cF: "solid looking rock."

The *Vista* response: Vista is scored when the subject's verbalization indicates that he has visualized the percept as being at an appreciable distance from himself. The critical element here is the spatial relationship between the subject's position in space in reference to the object perceived.

Rules for scoring vista/form dominance are analogous to those for color/form dominance. Examples are:

Free Assoc.		*Inquiry*			
II. D6	inside a cave, looking toward the entrance	shaped like a cave	Dsm	FV+	Ls
VI. D1	a bird's eye view of a mountain; the hills and valleys, looking down at it	from high up, an aerial view, rugged country, jagged outline	Dv	VF+	Ls
I. W	I get the impression of being up high	just the impression of depth	Wa	V	Abst

4. *The Perception of Activity*

"M" is scored when:

(*a*) Movement is perceived in human content. Movement includes, in addition to more overt activity:

(*i*) Postural references or any verbalized kinesthetic strain, e.g., "standing," "leaning," "bent," "twisted."

(*ii*) Part body postural referents, e.g., "pointing," "looking."

Frequently verbalizations of activity (primarily gestural) are associated with descriptions of inner states or functions, e.g., "angry," "brooding." Infrequently such descriptions are elicited without verbalization of manifest bodily activity and so are not scored M.

(*b*) Movement that is uniquely human is perceived in animal content, e.g., "dogs gossiping," "lions dancing."

"FM" is scored for those responses not scored as M in which movement is perceived in animal content.

"m" is scored for action in (*a*) inanimate objects or (*b*) for natural or abstract forces. Examples are:

(*a*) "volcano erupting," "top spinning," "rocks falling down"

(*b*) "something falling apart," "power radiating from the center," "something hanging from edge"

Not scored as "m" are:

(*a*) expressive descriptions of parts of living creatures, e.g., "grotesque faces" or "threatening images" or

(*b*) ambiguous dynamic terms, e.g., "stretched out," "spread out," "attached" . . .

5. *Adequate Organization of Blot Elements*

An "Integrative" scoring requires the appropriate organization (i.e., not fabulized combination, confabulation or contamination) of two or more adequately perceived blot elements (i.e., not F—) into a unified response. This includes the organization of:

(*a*) two or more "mediocre" elements

(*b*) two or more "vague" elements

(*c*) a "mediocre" and a "vague" element, or

(*d*) the combination of a "mediocre" or "vague" element with an "amorphous" element.

The elements in the responses may involve either identical or non-identical contents.

"Integration" is scored even if the basis for integration is developed only in the inquiry and modifies the original percept.

Free Assoc.	*Inquiry*	
III. 2 men	They're dancing together	Di M+ H P

"Integration" is scored if a new percept is developed in the inquiry for which a relationship is established to the "basic" percept. Neither the additional percept *per se* nor the determinants associated with the additional percept is scored.

Free Assoc.	*Inquiry*	
VIII. 2 animals	climbing a green mountain	Di FM+ A P

"Integration" is scored when a relationship is established between an adequate percept on the blot and an imaginal object not seen on the blot itself.

Free Assoc.	*Inquiry*			
IV. W A monster	looking for a victim	Di	M+	(H)

"Integration" is *not* scored when additional content (e.g., clothing) occupies all or part of the area used for the original percept.

Free Assoc.	*Inquiry*				
III. 2 people	They have black tuxedoes on, or	Dm	FC'+	H,Cg	P
	They're wearing high heel shoes	Dm	F+	H,Cg	P

The Functional Integrative Response: Functional integration is based upon an interaction between two or more conceptually independent subunits, i.e., movement of at least one object in relationship to at least one other object. Functional integrative is scored when, in addition to meeting the criterion of appropriate organization, the response contains M, FM, or m as a determinant. Examples:

III. D1 Two people holding something Di^{f} M+ H P

VIII. W Two beavers climbing a mountain Wi^{f} FM+ A P

VI. A plant (D3) coming up through soil (D1) Wi^{f} Fm+ Bt

Functional integrative responses are given weighted scores as follows:

(1) Active human content—score 4

A weighted score of 4 is given to a functional integrative response which is an M response in human content, and in which the movement is active. Examples:

2 figures dancing, men lifting up some object, a man riding horseback, etc.

Such verbs as fighting, walking, grasping, climbing, pushing, and the like would be included in the class of active verbs and thus would receive a weighted score of 4. (Where movement is not stated in verb form, the full weight is given to the appropriate noun, e.g., 2 dancers, 2 fighters, etc., given score of 4.)

(2) Active animal content—score 2

A weighted score of 2 is given to a functional integrative response which is FM, M in animal content, or M in anatomy content, where the movement is active. Examples:

FM: 2 bears climbing a hill, 2 dogs sniffing each other, 2 lobsters grasping something in their claws, etc.

M in animal content: 2 bears talking to each other, 2 dogs gossiping, 2 animals laughing at each other.

M in anatomy content: 2 skeletons dancing

Also given a score of 2 are such representations of action as "a scene from Night on Bald Mountain," "devil insects having a festival."

(3) Passive or static movement—score of 1

A weighted score of 1 is given to a functional integrative response where movement (either human or animal) is passive or static. Examples: Verbs connoting *passive receptive attitudes* such as drinking, sucking, warming hands over fire, cooking; *static movement* such as sitting, facing, looking, staring, holding, touching, standing; *stopped motion* verbs such as about to, trying to, going to, getting ready to, engaged in; *flector verbs* such as bending, leaning, sleeping, bending down, stooping, bowing.

General Rules for Quantitative Scoring

1. The first verb mentioned is the determinant of weighted score.

Example:	Two men facing each other, talking	Score 1
	Two men talking, facing each other	Score 4

2. Percepts must be complete figures (No Hd or Ad), for weighted scores to be given.

Example:	Two heads, looks like they're arguing.	Score 0
	Women looking at each other (inquiry: from the bust up)	Score 0

3. Inanimate movement is scored 0.

Example: Water flowing over rocks.	Score 0

"Collective" Integration: Here an overall grouping is actively imposed on the blot. The esential feature is evidence of a conceptualizing, classifying process as distinct from one of sheer enumeration. If the form-level of the class term is F+, an "integrative" scoring is given, even though subsequent enumeration may involve F— elements. Some of the key words which are indicative of "integration" are: group, garden, pair, and collection. Examples of some responses which are collective integrative are:

IV.	D6 —	pair of boots	Di^{c}	F+	Cg
VII.	W —	group of islands	Wi^{c}	F+	Ge
VIII.	W —	flower garden	Wi^{c}	F+	Bt
X.	W —	ocean scene	Wi^{c}	F+	Na
		collection of insects	Wi^{c}	F+	Na

Part of this category includes such key words as: bust of, emblem with, coat of arms with, ornament of, and statue of. Examples are:

II.	D1 —	emblem with 2 bears	Di^{c}	F+	Embl.
VII.	D1 —	statue of 2 dogs	Di^{c}	F+	A

Examples which do *not* reflect "collective" integration, but only numeration are:

"lots of . . . ," "a few . . . ," "couple of . . . ," "two . . . ," "picture of . . . ," "drawing of . . . ,". "Bunch" may or may not be integrative, depending on context, e.g., "bunch of kids" (no); "bunch of grapes" (yes).

"Positional" Integration: When there is a spatial relationship between the elements perceived, the response is scored "positional" integrative. Examples:

VI. W — totem pole set on a rock	Wi^{P}	F+	Ay	
	Dv	F+	Ls	
X. D11 — a limb with bugs on it	Di^{P}	F+	Bt	
	Dm	F+	A	
II. D6 — 2 dogs, nose to nose, barking	Di^{P}	FM+	A	P

"Structural" Integration: Here a single content is arrived at by unifying two or more colored areas, including the use of white space for its form characteristics. The essence of this category is color-boundary-crossing. Examples of W responses (Wi^{s}) are:

II — volcano; butterfly; clowns (even if not integrative based on M)
VIII — flower; fountain; boat
IX — fountain; gladiola bulb

Examples of structural integration involving D areas (Di^{s}) are:

II — lateral ½—man
VIII — D8—tree
IX — D9—mushroom from atom bomb

Note: a single content which falls into the "vague" category even though it crosses color boundaries continues to be scored "vague," e.g., Card VIII—W—design; Card IX—W—map; Card X—W—explosion.

The crossing of color boundaries is *not* "integrative" where one of the colors:

(*a*) occupies substantially all of the area required for a percept
(*b*) occupies only a relatively small area, and
(*c*) is not explicitly referred to by the subject

For example, D2 Card X is not "integrative" for "lion," but it is "integrative" for "fried egg" where the yolk (Dd33) is pointed out.

Integrative scoring can be assigned only once in any given response, regardless of the number of ways in which integration comes about. However, the numerical functional integrative score will always be given when it occurs.

6. *Inadequate Organization of Blot Elements*

"Fabulized Combination" is scored when *both* of the following criteria are met: (*a*) Responses in which two or more areas, separately interpreted, are combined on the basis of spatial relationship in a way which does not reasonably occur in nature, *and* (*b*) such a combination is not explained nor is it reasonably rationalized in a spontaneous fashion, i.e., without questioning by examiner.

Rules for "double scoring" apply to "fabulized combination." The "fabulized combination" score is given to the percept which occurs first in the free association. The "fabulized combination" score does not influence the scoring of form level of the individual percepts. Some examples of "fabulized combinations" are:

VII — W	Couple of baby lambs in the clouds dancing			
		Wfab	FM—	A
		Dv	F+	Cl
X — D12+D1	Couple of strange little green birds, standing on some organ of the body.			
		Dfab	FC.M—	A
		Dv	F+	An
V — W	Looks like a mule with wings. That's all.			
		Wfab	F—	Ad
		Dm	F+	Ad

"Fabulized combination" is not scored when there is adequate rationalization. Examples which are *not* "fabulized combinations" are:

I — W	Or it could be a person with butterfly wings in a dance recital	Wi	M+	H,Rc
		Dm	F+	Ad,Cg

Note: "In a dance recital" rationalizes the response.

VII — W	Some kind of nymphs dancing on the top of a cloud			
		Wi	M+	(H)
		Dv	cF.C'F+	Cl
	Inq. (D4) and nymphs is supposed to perform the supernatural.			

Note: The content "nymphs" serves to rationalize the response.

"Contamination" is scored when two independent responses to the *same* area are fused into a single percept.

Since "contamination" involves two percepts, the response is scored twice. The "contamination" score is given to the percept which occurs first in the free association. "Contamination" does not influence the scoring of form level of the independent percepts. Some examples of "contamination" are:

IV — W	Front of a bug—the front of an ox—the front of a bug—ox.	Wcontam	F—	A
		W—	F—	A

Note: F— is scored for both responses because both "bug" and "ox," if given independently, would warrant F— scoring.

Free Assoc.	*Inquiry*			
VIII. D5 Flag in the sky	Flag because of the form and sky because of the color	Dcontam Da	F— C	Obj Na
IX. D1 looks like grass, also looks like a rabbit—a grassy rabbit	grass because looks like tall grass and it's green shape of a rabbit	Dcontam D—	CF+ F—	Bt A

"Confabulation" (DW,DdW,DD,DdD) is scored when *both* of the following criteria are met:

(*a*) The interpretation of an area determines the interpretation of other areas, *and*

(*b*) The content developed is not a minor elaboration of the initiating percept. There are two types of "confabulated" responses. These are:

(*a*) response to a larger area which is determined by the response to a smaller one, *and*

(*b*) response in which the content given to a detail is given also to a larger blot area. "Confabulation" does not influence the scoring of form level (+ or —). Some examples of type (*a*) "confabulation" are:

Free Assoc.	*Inquiry*			
VI. W Cat	Because of the whiskers (Dd26)	DdW	F—	A
IV. W A human body — a female body	Woman's sexual organs (D3) and if so, then whole thing would pertain to a woman's body	DW	F—	An,Sex

An example of the (b) "confabulation" is:

Free Assoc.	*Inquiry*			
VI. W Bird	because of the top part (D3) is a bird	DW	F—	A

C. The Research Use of Developmental Indices

Not all the developmental indices described above have proven of equal value in our research studies. Some of the scores which have contributed most are the W— and D— responses. In general, the occurrence of the various developmental forms of the whole response are statistically unrelated to Rorschach response total. Empirically, we have found the D— per cent score, $\frac{(D—)}{(D\quad)}$, which is also independent of total number of responses, a particularly valuable index. Among the determinants, the M response has shown the highest number of relationships followed in turn by the FC and the absolute number of C plus CF (C+CF) responses. We have also developed a number of composite scores which have proved of value. These are: (*a*) the functional integrative score (Fi) described above, (*b*), an X+XF score which is equal to the absolute total of C+CF, C′+C′F, c+cF, and V+VF responses, (*c*) an analogous FX score, and (*d*) an index of primitive thought (IPT), composed of the absolute number of fabulized combinations, contaminations, and confabulation responses.

The frequency of many of these indices in any individual Rorschach record is often related to the total number of responses (R). However, transformation to percentages of R is often an inappropriate control in that it rests upon untenable assumptions regarding the nature of the relationship between R and the given index. Further, the magnitude of such transformed values is often more directly reflective of variations in R than of differences in frequency of occurrence of a given response-type.

Accordingly, when dealing with indices that are related to R, we have formed "high" and "low" R groups by dichotomizing the sample at the median of the R distribution. Statistical comparisons involving frequencies of response in the genetic categories have then been made separately within each R group. To arrive at inferences relevant to the total sample we have then combined the separate results by Fisher's method for the combination of tests of significance.

In general, the use of distribution-free procedures seems indicated for analyses involving these developmental indices. Most of the distributions are markedly skewed. Further, the developmental "scores" can logically be interpreted only as reflecting an ordering of individuals with respect to the presumed underlying psychological dimensions. In practice, we have typically reduced the number of positions in the ordinal scale for any given index to any two, e.g., absence and presence. This procedure is based on an analysis of the nature of the Rorschach test situation and its reasonableness is confirmed by certain empirical findings. Thus, if a subject develops a given type of response, we have direct evidence that he is capable of so doing; but if he fails to do so, it is probably inappropriate to infer that the corresponding psychological function is necessarily lacking. Many factors may contribute to its absence from the protocol, e.g., anxiety level, the subject's implicit assumptions about the nature of the test situation, his motivations, etc. Empirical support for this consideration stems from an examination of the relationships between the developmental indices and various non-Rorschach variables (e.g., performance under stress). We have often noted that among those subjects who do not develop a single response appropriate for a given index there is no systematic tendency for them to rank high (or low) on the criterion measure. On the other hand, those subjects who give at least one response in the category are relatively homogeneous with respect to ranking (either high or low) on the same non-Rorschach variable.

As stated above, a subject may be distracted from the production of a given type of Rorschach index for a variety of reasons. Consequently, it may be inappropriate to conclude, for example, that the occurrence of three responses of a given class reliably reflects greater strength of the presumed psychological function than that indicated by two responses of that type. Consequently, in our analysis of the developmental indices we have mainly employed simple two-point scales, e.g., "few" versus "many," "absent" or "present," etc.

REFERENCES

1. BECK, S. J. Rorschach's Test: Volume 1, Basic Processes (Rev. Ed.) New York: Grune & Stratton, 1949.
2. BECKER, W. C. A genetic approach to the interpretation and evaluation of the process reactive distinction in schizophrenia. *J. Abn. & Soc. Psychol.*, 1956, **53**, 229-236.
3. DWORETZKI, G. Le test de Rorschach et l'evolution de la perception. *Arch. de Psychol.*, 1939, **27**, 233-396.
4. FORD, M. The application of the Rorschach test to young children. Univ. Minnesota Institute of Child Welfare Monographs, No. 23, 1946.
5. FOWLER, R. Psychopathology and social adequacy: A Rorschach developmental study. (Unpublished doctor's dissertation. Pennsylvania State Univ., 1957.)
6. FRAMO, J. L. Structural aspects of perceptual development in normal adults: a tachistoscopic study with the Rorschach technique. (Unpublished doctor's dissertation, Univ. Texas, 1952.)
7. FRANK, I. Perceptual structurization in certain psychoneurotic disorders: a genetic evaluation by means of the Rorschach test. (Unpublished doctor's dissertation, Boston Univ., 1951.)
8. FREED, E. Perceptual differentiation in schizophrenia: a tachistoscopic study of structural Rorschach elements. (Unpublished doctor's dissertation, Syracuse Univ., 1952.)
9. FRIEDMAN, H. Perceptual regression in schizophrenia: An hypothesis suggested by the use of the Rorschach test. *J. Genet. Psychol.*, 1952, **81**, 63-98.
10. ———. Perceptual regression in schizophrenia: An hypothesis suggested by the use of the Rorschach test. *J. Proj. Tech.*, 1953, **17**, 171-185.
11. GOLDSTEIN, K. The Organism. New York: American Book, 1939.
12. GRACE, N. B. A developmental comparison of word usage with structural aspects of perception and social adjustment. (Unpublished doctor's dissertation, Duke Univ., 1956.)
13. HEMMENDINGER, L. Perceptual organization and development as reflected in the structure of the Rorschach test responses. *J. Proj. Tech.*, 1953, **17**, 162-170.
14. HERSCH, C. Perceptual structure in creative artists: An analysis by means of the Rorschach test. (Unpublished doctor's dissertation, Clark Univ., 1957.)
15. HURWITZ, I. A developmental study of the relationship between motor activity and perceptual processes as measured by the Rorschach test. (Unpublished doctor's dissertation, Clark Univ., 1954.)

16. KLOPFER, B., MARGULIES, H., MURPHY, L. B., & STONE, L. J. Rorschach reactions in early childhood. *Ror. Res. Ex.*, 1941, **5**, 1-23.

17. KRUGER, A. Direct and substitute modes of tension-reduction in terms of developmental level: An experimental analysis of the Rorschach test. (Unpublished doctor's dissertation, Clark Univ., 1954.)

18. LANE, J. E. Social effectiveness and developmental level. *J. Pers.*, 1955, **23**, 274-284.

19. LOFCHIE, S. H. The performance of adults under distraction stress: A developmental approach. *J. of Psychol.*, 1955, **39**, 109-116.

20. MISCH, R. C. The relationship of motoric inhibition to developmental level and ideational functioning: An analysis by means of the Rorschach test. (Unpublished doctor's dissertation, Clark Univ., 1954.)

21. MOREAULT, L. Le test de Rorschach et le fonctionnement intellectual dans la schizophrenie benigne. (Unpublished doctor's dissertation, University of Montreal, 1956.)

22. MURPHY, G. Personality. New York: Harper, 1947.

23. PEÑA, C. D. A genetic evaluation of perceptual structurization in cerebral pathology: An investigation by means of the Rorschach test. *J. Proj. Tech.*, 1953, **17**, 186-199.

24. PHILLIPS, L., & COWITZ, B. Social attainment and reactions to stress. *J. Pers.*, 1953, **22**, 270-283.

25. PHILLIPS, L., & SMITH, J. G. Rorschach Interpretation: Advanced Technique. New York: Grune & Stratton, 1953.

26. PIAGET, J. The Psychology of Intelligence. New York: Harcourt, Brace, 1950.

27. RICCIUTI, H. N. Use of the Rorschach test in longitudinal studies of personality development. *J. Proj. Tech.*, 1956, **20**, 256-260.

28. ROCHWARG, H. Changes in the structural aspects of perception in the aged: An analysis by means of the Rorschach test. (Unpublished doctor's dissertation, Michigan State Univ., 1954.)

29. ROSENBLATT, B., & SOLOMON, P. Structural and genetic aspects of Rorschach responses in mental deficiency. *J. Proj. Tech.*, 1954, **18**, 496-506.

30. SANDER, F. Experimentelle Ergebnisse der Gestalt Psychologie. *In* E. Becher *(Ed.)*, *Bericht uber den 10, Kongress fur experimentelle Psychologies,* 1927. Jena: Fischer, 1928. (Pp. 23-88.)

31. SIEGEL, E. L. Genetic parallels of perceptual structurization in paranoid schizophrenia: An analysis by means of the Rorschach technique. *J. Proj. Tech.*, 1953, **17**, 151-161.

32. SULLIVAN, H. S. The Interpersonal Theory of Psychiatry. New York: Norton, 1953.

33. WALDMAN, M. Personality factors and performance under stress. (Unpublished doctor's dissertation, Univ. Chicago, 1956.)

34. WERNER, H. Comparative Psychology of Mental Development. (Rev. Ed.) Chicago: Follett, 1948.

35. ———. The concept of development from a comparative and organismic point of view. *In* D. B. Harris (*Ed.*), *The Concept of Development: An Issue in the Study of Human Behavior.* Minneapolis: Univ. Minnesota Press, 1957.

36. WILSON, M. T. Regression in perceptual organization: A study of adolescent performance on the Rorschach test. (Unpublished doctor's dissertation, Clark Univ., 1954.)

Worcester State Hospital
Worcester 1, Massachusetts

The Journal of Genetic Psychology, 1959, **94**, 287-296.

THE FUNCTION OF VERBAL LABELS IN THE DISCRIMINATION OF SUBTLE STIMULUS DIFFERENCES*

Massachusetts Mental Health Center, Boston, Massachusetts

ROBERT S. ALBERT[1]

A. INTRODUCTION

It is commonly held that association of a verbal label to a non-verbal stimulus object automatically enhances that object's discriminability; and, hence, *S*'s control over the object (1, 2, 4, 10, 21). Part of the same hypothesis holds such labelling one factor making social symbols and stereotypes so effective as guides and controls in social behavior. Occasions where this hypothesis has been demonstrated empirically were either in transfer studies (2, 4, 15, 17, 22) or perceptual discrimination (5, 10, 14, 18, 19, 20, 27). However, most often the stimuli were non-meaningful stimuli of obvious discriminable differences, meaningful or highly representative stimuli. The present study also is concerned with testing this conception of labels but with non-meaningful stimuli of empirically determined non-discriminableness. The hypothesis tested is whether or not associated verbal labels make possible a discrimination where previously *S* was unable. Stated specifically, the general hypothesis under examination is that verbal labels, functioning as cue-producing responses (7), facilitate the discrimination of empirically determined non-discriminable stimuli. More specific aspects underlying this hypothesis result in the following three experimental hypotheses:

1. The greater *S*'s previous experience in using linguistic labels, the greater the enhancement of his ability to discriminate when using labels paired with non-meaningful stimuli, i.e., adults, having had more experience than children in labelling and more verbal associations with which to associate, will have more changes than children in the direction of making more correct labelled discriminations.

2. An inverse relationship exists between "non-verbal" ability to discriminate and the function of linguistic cues. Where *S* is less able to make

---

* Received in the Editorial Office on March 28, 1958.

[1] The writer would like to thank Dr. A. J. Brodbeck for commenting upon a much earlier draft of the paper. His and Dr. William Bevan Jr.'s critical and constructive appraisals are greatly appreciated. Thanks are also due to Mr. Aron Saachs and associates of Hecht House, Dorchester, Massachusetts, for making available the young subjects and research facilities.

the correct non-labelled discrimination, the influence of the labels will be such as to make the change in the label-mediated performance of "poor" and "good" non-verbal discriminators not statistically significant. Labels compensate for poorer non-verbal discriminating ability.

3. The greater the number of learning trials in which *S* is presented both stimuli and appropriate verbal labels, the more correct *S*s' discriminations will become. The more "pooled" labelled experience *S* has to draw upon, the sharper the stimulus differences will appear; *S* will have a clearer perception of the empirical referent, i.e., the meaning of the verbal label (this will be discussed more fully below).

These three hypotheses are meant to take into consideration the salient characteristics of the sign-referent relationship, namely: (*a*) *induction* is the process by which the relationship between sign and referent is acquired, (*b*) signs stand in "close" temporal and/or physical contiguity to the referent, and (*c*) the particular aspect of the stimulus experience for which the sign is to stand must be invariant in terms of its position along some dimension, i.e., the word 'lighter' signifying only those empirical differences of a critical (judged) stimulus weighing physically less than the compared stimulus.

B. Method

1. *Subjects*

Twenty-five children, seven to 10 years old (mean age: 8 years, 9 months) (17 girls, eight boys). Thirteen college, adult subjects (seven females, six males).

2. *Stimulus Material*

Six metal cylinders of uniform size containing lead *BB* weights; cotton enveloped the weights to eliminate auditory cues. The cylinders were kept covered when not used to eliminate temperature differences between used and unused weights. The weights were, in grams: 19.60, 20.16, 20.72, 21.28, 21.84, and 22.40.

3. *Procedure*

The experiment was divided into two parts. In part one *S*s were operationally divided into two groups: "good" or "poor" discriminators of non-labelled weights. Adult and young *S*s were categorized separately. Part one was presented to the young *S*s as a game and to adults as an experiment in psychophysical measurement. Post-experimental questioning indicated that *S*s accepted the stated purpose.

*S*s' Part One discrimination scores were determined by adding together the total number of errors within the experimental sequence, *plus* the empirical weight difference of the cylinders between which errors were made. "Good" signifies *low* error-scores; "poor," *high* error scores.

*S*s used in Part Two could be divided operationally into two statistically significant different groups of "good" and "poor" non-verbal discriminators. The differences in discrimination scores for these two categories among both adults and children had P-values $< .01$. Interestingly enough, the difference between adult and young "poor" discriminators in Part One was also statistically significant ($P < .05$), young *S*s performing better than the adults. There is no statistically significant difference between adult and young "good" discriminators ($P < .9 > .8$).

Adults were presented the cylinders in the following order[2]: 5, 8, 7, 7, 6, 9, 10, 7, 9, 8, 7, 7, 6, 5, 5, 6. After the first stimulus was presented adult *S*s were asked to judge if succeeding stimulus was "heavier," "lighter," *or* the "same" as the preceding stimulus.

Young *S*s were presented the cylinders in a 5, 8, 7, 6, 9, 10, 7, 9, 8, 7, 6, 5, 6 order, and were asked to judge if the succeeding stimulus was "heavier" *or* "lighter." The response "same" was not asked since pre-testing showed that young *S*s habitually gave this verbal response regardless of the actual difference between the stimuli.

Part Two was carried out seven to 10 days later on a total of 24 *S*s (12 children, 12 adults; each subgroup was made up of an equal proportion of males and females, "good" and "poor" discriminators).

In Part Two *S*s closed their eyes and were presented four cylinders between which there were no *JND*'s (28). These cylinders were presented one at a time. *E* verbalized the appropriate label; *S* repeated the label after *E* had said it and had placed the corresponding weight in *S*'s preferred hand for 10 seconds.

The names given the four cylinders were "Bee," "Bell," "Ball," and "Bang" respectively. *S*s were given either 1, 3, or 5 presentations. Afterward *S*s were shown four randomly placed cardboard squares on a table. On each square was printed one of the four labels. All *S*s were then presented the four critical stimuli one at a time, in a 5, 6, 7, 8, 8, 6, 7, 5 order, and asked to place each on the appropriately labelled square.

---

[2] Number equals the number of lead BB shots in a cylinder. The numbers were retained to designate the cylinders.

C. RESULTS

Results show that differences in labelled discrimination performances between "good" and "poor" discriminators are not statistically significant for adults or children (P-values $< .7 > .5$, .99 respectively). Comparing "good" adults with "good" young Ss, and "poor" adults with "poor" young Ss in labelled discrimination performances, the data show almost statistically significant differences in the opposite to predicted direction in each case (Chi-square $= 2.262$, $df = 1$, 1-tail $P < .1 > .05$; chi-square $= 1.942$, $df = 1$, 1-tail $P < .1 > .05$). In both cases young Ss perform better. Furthermore, comparing *all* adult Ss with all young Ss shows the young Ss' performances with labelled stimuli the better of the two groups (chi-square $= 5.228$, $df = 1$, 1-tail $P < .025 > .01$).

However, the predicted greater enhancement for adults' labelled stimuli discriminations was confirmed. The Sign test (26) was used, having $+ =$ a change from an equal or a greater number of errors than correct responses in Part 1 to more correct responses than errors in Part 2. For all adults there were 7 changes, all positive ($P = .016$); for young Ss there were 9 changes, 8 positive ($P = .04$).

The prediction of the second hypothesis was also confirmed. The positive changes in the labelled discriminations of "good" and "poor" Ss (whether among adults or children) were not statistically significant. Among adults the frequencies were 4 +'s, 2 no-changes; 3 +'s, 3 no-changes for "good" and "poor" discriminators respectively. Among young Ss comparable subgroup frequencies were 5 +'s, 1 no-change; 4 +'s; 2 no-change or poorer performances. None of the Fisher Exact Test P-values were statistically significant.

It also seems that for adults previous labelling experience may be a handicap within a new situation. Post-experimental questioning of the adults showed that they often tried to formulate verbally the stimulus differences in Part 1 and find some rule by which to direct their performance throughout the sequence. That this attempt to formulate a rule leads to a poorer adult performance agrees with other data (19, 25) which also show that too early an attempt to verbalize perceptions leads to a poorer mastery of the concept material than a later attempt after the non-verbal material or conceptual principle has been better assimilated. Apparently covert integration must take place before labels can acquire a substantive meaning. Too early an attempt distracts Ss' attention and may result in the wrong attribute, etc., being labelled.

It should be pointed out that the number of correct responses for all *S*s in Parts 1 and 2 is less than expected by chance. This indicates to the writer two things: (*a*) that the empirical differences among the stimuli were too small to be easily perceived; (*b*) more important, that the presence of simple linguistic labels functioning as cues did *not* counteract the "raw" indiscriminability among the stimuli, increasing their discriminability. The general notion holding for labels needs amending for perceptible differences apparently must be initially discriminable before one can use linguistic labels to positive advantage. This seems to hold true within this experiment regardless of *S*s' non-verbal perceptual abilities, age, or sex.

TABLE 1
INDEPENDENCE BETWEEN NUMBER OF LEARNING TRIALS AND INFLUENCE OF LINGUISTIC CUES IN DISCRIMINATION

	Comparison	Chi-Sq.	df	1-tail P-value	No. of Trials	Freq. of Part 2 Responses	
						Right	Wrong
Adults	*S*s with 5 trials vs. Ss with 3 trials	.061	1	.9	5)	11	21
	*S*s with 5 trials vs. Ss with 1 trial	3.233	1	<.05>.02	3)	12	20
	*S*s with 3 trials vs. Ss with 1 trial	4.082	1	<.03>.01	1)	4	28
Children	*S*s with 5 trials vs. Ss with 3 trials	.564	1	<.25>.15	5)	13	19
	Ss with 5 trials vs. Ss with 1 trial	.271	1	<.35>.25	3)	17	15
	*S*s with 3 trials vs. Ss with 1 trial	2.306	1	<.1 >.05	1)	10	22
	Adults and child. each group with 5 trials	.250	1	<.7 >.5*			
	Adults and child. each group with 3 trials	1.008	1	.4*			
	Adults and child. each group with 1 trial	2.285	1	<.2 >.1*			

* 2-tail *P*-values (8).

Data in Table 1 partially confirms the third hypothesis, i.e., that the more opportunities one has for associating a verbal label with its empirical referent, the greater the number of subsequent correct perceptions of that referent. However, this conclusion seems more applicable to adults' than children's performance. The Part 2 performance differences were statistically significant in two of the comparisons among adult *S*s; but approached statistical significance only once among the performances of children.

When different degrees of non-verbal discriminative ability are randomized among all *S*s and the performance as a function of number of learning trials is compared between adult and children, no statistically significant differences appear between the two groups.

D. General Discussion

Research in this area seems prone to raising more questions than giving answers. The materials dealt with, the processes investigated, are illusive; if anything, they are self-concealing and self-cancelling. Verbalization obscures the very processes one wishes to study. However, there seem several conclusions one can draw from the study. The first concerns the general theoretical hypothesis under examination, namely that attached verbal labels, as cue-producing responses, facilitate discrimination among otherwise non-discriminable differences. The hypothesis received very limited verification. It seems such a rôle for verbal labels is limited to material that may be initially discriminable although perhaps *not* identifiable by *S*, a distinction similar to that between the perception of unity within figure-ground relationships and of identity (12) which need not occur simultaneously within *S*'s genetic development or experience. Figure-ground perception is a fairly autochthonous process; the perception of identity is learned—a matter of much experience and generalization. Such a distinction is also similar to the distinction between the non-verbal unconscious and the "verbalizable" preconscious images posited by psychoanalysis. In the case of the latter awareness is partial and operative; in the case of the former past experience is operative but not the object of awareness or identifying processes—without outside help.

The second general conclusion is more a matter of inference from the performance and reports of the adult *S*s. Verbal labels rather than being simply "cue-producing responses" seemed to function more as "cue-producing *stimuli*." This function may confound, not simplify, the underlying process. Rather than one's responses being directly governed (cued) by the associated verbal meaning, upon hearing a label one is stimulated by this verbal cue to *think* about, or recall, the many associated empirical references to which the verbal cue refers, after which a decision is made concerning *S*s' responses. Associating the correct referent to the label is far from automatic, but an instance of cognition, especially when one must supply the label (1). This means that elaborate covert processes of recall, comparisons, and decision-making precede the "final cued" response. Such complexity may be particularly evident in the early stages of verbal learning and/or when the empirical referents are not easily distinguished, e.g., learning to recognize new model cars, identical twins, signs of weather changes, etc.

Concerning this importance of stimulus differences among fairly similar stimulus objects, Bevan and Saugstad (3) found that to achieve preciseness and greater uniformity of judgments among comparably trained *S*s, a wide range of stimuli with a wide inter-stimulus difference is needed. Where the range is short and the inter-stimulus difference slight, as in the case of the present experiment, the accuracy of *S*s' judgments suffers considerably. Data reported by Crossman (6) confirms this. In experiments in the various perceptual modalities he, too, found that the greater the similarity of adjacent stimuli (which is analogous to a narrow inter-stimulus difference), the greater the difficulty in discrimination; pari passu, the greater *S*'s "confusion" between stimuli. Stimuli, it must be remembered, often have built into their sequential occurrence physical relationships to and among one another which obscure and even preclude their discrimination as different. It would be interesting (and important) to find that range of stimuli and its combined width of inter-stimulus difference that labelling is most facilitative in aiding the making and retaining of discriminations. It strikes the writer as being just above the *JND* in which case one would have enough of an empirical difference not only to perceive, but be referable to, i.e., identifiable.

It should be kept in mind that there are two important differences between these studies and the present one. In (3, 6) subjects were presented with a constant difference between stimuli. In the present study stimulus differences shifted within Part 1 although not within Part 2. This may influence the selection of "good" and "poor" discriminators for Part 2 by introducing some personality variable such as flexibility in shifting set, which may differ from, and be additional to, the ability to discriminate non-verbally. The other difference between the studies is in the subjects used. Here adults and children were used; the other studies used only adults. Their results help explain the adult data in the present study. However, it is still moot how applicable their data are in explaining children's responses in the present study.

Another point which needs some discussion is the effect the number of trials had upon *S*s' use of labels, especially for the adults. It seems that as *S* acquires additional experience with the label *and* its empirical referent, such increased experience (number of trials within this experiment) would result in a shift of the mid-point within *S*s' adaptation level for that particular kind of stimulus material (1, 13). Such a shift would mean an extension of the range *and* the anchor points because of the insertion of an additional difference between the two points. This could have two effects on the conditions underlying *S*s' discriminations: to widen the difference among stimuli and increase their discriminability, and to clarify the meaning, i.e., make more referable, "isolate," the empirical referents of the associated linguistic labels.

Two additional data remain to be discussed—the initial superiority of children over adults in making non-verbally mediated discriminations and adults' improved performance with more learning trials. One questions if children's initial superiority would hold up over time, but may not be a function of the initial "complicating" adult attempts to formulate the stimulus differences and also the presence of different cognitive sets operating in the performances of the two groups. The two influences are possible; and the possibility of adult performances becoming better with time and being maintained longer exists.

Riess (23) has shown that different semantic bases may effectively mediate transferred conditioned responses. However, a different semantic basis predominates at a different chronological age. There were two groups of *S*s within his experiment—one between seven and nine years, the other between 17 and 20 years old—which fit closely the two present experimental groups. Riess found that the predominant semantic basis for the young and adult *S*s within his study differed; homophones for the young, synonyms for the adults. This difference points up the possibility that certain developmental changes may have as correlates changes in the relationships between linguistic variables and their referents. While it is difficult to say that there is an empirical analogue for homophones, equally superficial and obvious similarities *can* be found and used in behavior. The relationship of synonymity is somewhat more subtle; it is contingent upon *S* perceiving similarity between or among things which ostensibly need not, or do not, show such similarity. It is relational, abstract; not concrete or perceptual.

It seems reasonable to assume that the more experience one has, the larger number of different aspects of each stimulus one may become familiar with and the more abstract and numerous the similarities possible among the stimulus objects. Where experience is limited with new stimuli, gross categorization would result, thereby limiting the possibilities of similarities to the more obvious.

Riess' data fit Garrett's (9) developmental theory of intelligence. Cognitive processes conventionally termed *intelligence* change from being highly unified and amorphous to specific abilities, more or less separate in their function and increasingly independent of one another. Garrett states that "with increasing age there appears to be a gradual breakdown of the amorphous general ability into a group of fairly distinct aptitudes. It seems highly probable that *maturation* has much to do with this differentiating process, but *increasing experience* and diverging interests must also contribute heavily," (Italics mine). He reports inter-correlations between *S*s' performances on verbal and non-verbal tests that consistently decrease as *S*s' ages increased. This change appears analogous to the increased specification of the verbal processes found by Riess.

It would appear that as one ages and has greater experience with the world about him, one also can categorize its manifold stimulus features along an increasing *number* of relevant dimensions; these responses are learned first hand or vicariously through culture (21); their relevancy increases along many different dimensions as we move from restricted egocentric contact (21, 24). As a summary of these changes one can presume that the adult differs from the child in two important respects. He has acquired and uses more and finer scaled dimensions for his increased "kinds" of experiences; secondly, the dimensions' ranges widen, i.e., they contain more discriminable intervals and "wider" anchor points. The child's may be termed more "broad," in the sense of being more amorphous and inclusive, containing fewer internal difference points,[3] e.g., *all* sweets are good, *all* dogs fear-provoking, etc.

Learning to perceive *identity* seems to be a function of generalizing one's experience. The transfer or generalization of discriminations implies that abstraction has occurred and underlies the transfer (10). Abstraction can be carried out on *N* dimensions. The only requirement is that *S* be aware of a dimension common to two or more objects. One such basis for abstraction is "meaning." Riess' experiment shows that a shift occurs with age in the semantic basis for transfer of a response. It may be a case of *S*s of different ages being able to put different questions to their experiences (24). The adult—especially the college trained—asks what differences and similarities occur and are common to objects, events, etc., within his experience; the young child generally initially assumes the external world. Postulating such questioning and underlying processes may help to account for the adult's greater profit from learning trials with labels in the sense that they come to increasingly "mean," be associated, with something, whereas such trials did not seem to effect so quickly children's use of labels in identifying different stimuli.

---

[3] Such an explanation in a sense accounts for the regular findings of overlap between age and intelligence in cognitive behavior. An increase in either can partially compensate for a lack in the other (16, 24).

REFERENCES

1. ALBERT, R. S. A discussion of the effects of signs and symbols upon behavior with an emphasis on linguistic variables. (To be published.)
2. ALBERTS, E., & EHRENFREUND, D. Transposition in children as a function of age. *J. Exp. Psychol.,* 1951, **41**, 30-38.
3. BEVAN, W., & SAUGSTAD, P. Breadth of experience, ease of discrimination, and efficiency of generalization. *Brit. J. Psychol.,* 1955, **56**, 13-19.
4. BINGE, J. S. Verbal responses in transfer. (Unpublished Ph.D. Dissert., Yale Univ., New Haven, 1941.)
5. BROWN, R. W., & LENNEBERG, E. H. A study in language and cognition. *J. Abn. & Soc. Psychol.,* 1954, **49**, 454-462.
6. CROSSMAN, E. R. W. The measurement of discriminability. *Quart. J. Exp. Psychol.,* 1955, **7**, 176-195.
7. DOLLARD, J., & MILLER, N. E. Personality and psychotherapy. New York: McGraw-Hill, 1950.
8. EDWARDS, A. L. Experimental design in psychological research. New York: Rinehart, 1950.
9. GARRETT, H. E. A developmental theory of intelligence. *In* (E. L. Hartley, *Ed.*) *Outside Readings in Psychology.* New York: Crowell, 1950.
10. GIBSON, E. J. Improvement in perceptual judgments as a function of controlled practice or training. *Psychol. Bull.,* 1953, **50**, 401-431.
11. HAZLETT, C. Children's thinking. *Brit. J. Psychol.,* 1930, **20**, 354-361.
12. HEBB, D. O. Organization of Behavior. New York: Wiley, 1949.
13. HELSON, H. Perception. *In* (H. Helson, *Ed.*) *Theoretical Foundations of Psychology.* New York: Von Nostrand, 1951.
14. HENDRIX, G. A new clue to transfer of learning. *Elem. Sch. J.,* 1947, **48**, 197-208.
15. JEFFREY, W. E. The effects of verbal and nonverbal responses in mediating an instrumental act. *J. Exp. Psychol.,* 1953, **45**, 327-333.
16. KRECH, D., & CALVIN, A. Level of perceptual organization and cognition. *J. Abn. & Soc. Psychol.,* 1953, **48**, 394-400.
17. KUENNE, M. R. Experimental investigation of the relation of language to transposition behavior in young children. *J. Exp. Psychol.,* 1946, **36**, 471-490.
18. KURTZ, K. H., & HOWLAND, C. I. The effect of verbalization during observation of stimulus objects upon accuracy of recognition and recall. *Amer. Psychol.,* 1952, **7**, 278.
19. KURTZ, K. H. Discrimination of complex stimuli: The relationship of training and test stimuli in transfer of discrimination. *J. Exp. Psychol.,* 1955, **50**, 283-292.
20. NORCROSS, K. J., & SPIKER, C. C. The effects of type of stimulus pre-training on discrimination performance in preschool children. *Child Devel.,* 1957, **28**, 79-83.
21. PIAGET, J. The Psychology of Intelligence. New York: Harcourt-Brace, 1950.
22. PYLES, M. K. Verbalization as a factor in learning. *Child Devel.,* 1932, **3**, 108-113.
23. RIESS, B. F. Genetic changes in semantic conditioning. *J. Exp. Psychol.,* 1946, **36**, 143-152.
24. SCHUESSLER, K. F., & STRAUS, A. L. Sequential stages in the understanding of the concept of money. *Amer. Sociol. Rev.,* 1950, **15**, 752-762.
25. SCHWARTZ, B. D. Certain relations between verbalization and concept formation. (Unpublished Ph.D. Dissert., Princeton Univ., 1948.)
26. SIEGEL, S. Non Parametric Statistics for the Behavioral Sciences. New York: McGraw-Hill, 1956.
27. SIGEL, I. E. The dominance of meaning. *J. Genet. Psychol.,* 1954, **85**, 201-207.
28. WOODWORTH, R. S. Experimental Psychology. New York: Holt, 1938.

Massachusetts Mental Health Center
74 Fenwood Road
Boston 15, Massachusetts

The Journal of Genetic Psychology, 1959, **94**, 297-300.

BOOKS

Now that there is a special *APA* journal completely devoted to the publication of book reviews, it is no longer necessary that other journals emphasize such publication. It has always been our conviction that book reviews are a secondary order of publication unless they carry information that is equally important as the book. However, the publication of book titles is a very important service, and we shall continue to render that service.

In any given issue of this journal, we may continue to publish one or more book reviews, but we do not consider such publication a major function of this journal. In line with this policy, we can no longer pay for such manuscripts.

BOOKS RECENTLY PUBLISHED

ABRAHAM, W. Common Sense About Gifted Children. New York: Harper, 1958. Pp. 268.

AINSWORTH, M. D., & AINSWORTH, L. H. Measuring Security in Personal Adjustment. Toronto: Univ. Toronto Press, 1958. Pp. 98.

ALLEN, R. M. Personality Assessment Procedures: Psychometric, Projective, and Other Approaches. New York: Harper, 1958. Pp. 541.

ANDERSON, C. C. Function Fluctuation. New York: Cambridge Univ. Press, 1958. Pp. 104.

ARGYRIS, C. Personality and Organization: The Conflict Between System and the Individual. New York: Harper, 1957. Pp. 291.

ATKINSON, J. W., *Ed.* Motives in Fantasy, Action, and Society: A Method of Assessment and Study. Princeton: Van Nostrand, 1958. Pp. 873.

AUSUBEL, D. P. Drug Addiction: Physiological, Psychological, and Sociological Aspects. New York: Random House, 1958. Pp. 126.

BAKAN, D. Sigmund Freud and the Jewish Mystical Tradition. Princeton: Van Nostrand, 1958. Pp. 326.

BANFIELD, N. C. The Moral Basis of a Backward Society. Glencoe: Free Press, 1958. Pp. 204.

BANKS, F. Teach Them to Live. New York: International Univ. Press, 1958. Pp. 287.

BARTLETT, F. Thinking: An Experimental and Social Study. New York: Basic Books, 1958. Pp. 203.

BEARDSLEE, D. C., & WERTHEIMER, M., *Eds.* Readings in Perception. Princeton: Van Nostrand, 1958. Pp. 751.

BERGLER, E. Counterfeit-Sex. New York: Grune & Stratton, 1958. Pp. 380.

BOSS, M. The Analysis of Dreams. New York: Philosophical, 1958. Pp. 223.

BROWN, C. C., & SAUCER, R. T. Electronic Instrumentation For the Behavioral Sciences. Springfield: Thomas, 1958. Pp. 160.

BROWN, H., *Ed.* Science and the Creative Spirit. Toronto: Univ. Toronto Press, 1958. Pp. 165.

BURTON, G. Personal, Impersonal, and Interpersonal Relations. New York: Springer, 1958. Pp. 230.

BYKOV, K. M. The Cerebral Cortex and the Internal Organs. New York: Chemical Publishing, 1957. Pp. 448.

CAUDILL, W. The Psychiatric Hospital as a Small Society. Cambridge: Harvard Univ. Press, 1958. Pp. 406.

CLARK, W. H. The Psychology of Religion. New York: Macmillan, 1958. Pp. 485.

COELHO, G. V. Changing Images of America. Glencoe: Free Press, 1958. Pp. 145.

COLE, L. E., & BRUCE, W. F. Educational Psychology. Yonkers-on-Hudson: World Book, 1958. Pp. 701.

CORNER, G. W. Anatomist at Large. New York: Basic Books, 1958. Pp. 215.

Deese, J. The Psychology of Learning. New York: McGraw-Hill, 1958. Pp. 367.

Durkheim, E. Professional Ethics and Civic Morals. Glencoe: Free Press, 1958. Pp. 228.

Edwards, A. L. Statistical Analysis. New York: Rinehart, 1958. Pp. 234.

English, N. B., & English, A. C. A Comprehensive Dictionary of Psychological and Psychoanalytical Terms. New York: Longmans, Green, 1958. Pp. 594.

Ekstein, R., & Wallerstein, R. S. The Teaching and Learning of Psychotherapy. New York: Basic Books, 1958. Pp. 334.

Emery, F. E., & Oeser, O. A. Information, Decision, and Action. Melbourne: Melbourne Univ. Press, 1958. Pp. 132.

Esman, A. H., *Ed.* New Frontiers in Child Guidance. New York: International Univ. Press, 1958. Pp. 218.

Fein, R. Economics of Mental Illness. New York: Basic Books, 1958. Pp. 164.

Fisher, S., & Cleveland, S. E. Body Image and Personality. Princeton: Van Nostrand, 1958. Pp. 420.

Frank, W. The Rediscovery of Man. New York: Braziller, 1958. Pp. 491.

Galdston, I. The Family in Contemporary Society. New York: International Univ. Press, 1958. Pp. 147.

Galdston, I., *Ed.* Panic and Morale. New York: International Univ. Press, 1958. Pp. 340.

Garma, A. Peptic Ulcer and Psychoanalysis. Baltimore: Williams & Wilkins, 1958. Pp. 143.

Garrett, H. E. Statistics in Psychology and Education. New York: Longmans, Green, 1958. Pp. 478.

Gee, H. H., & Glaser, R. J., *Eds.* The Ecology of the Medical Student. Evanston: Association of American Medical Colleges, 1958. Pp. 262.

Glaser, R., Zubin, J., Lindsley, D. B., Nowlis, V., Coffer, C. N., Guetzkow, H., Carroll, J. B., Hamlin, R. M., & French, T. M. Current Trends in the Description and Analysis of Behavior. Pittsburgh: Univ. Pittsburgh Press, 1958. Pp. 242.

Goldberg, S. Introduction to Difference Equations. New York: Wiley, 1958. Pp. 260.

Gorman, M. The Educational Implications of the Theory of Meaning and Symbolism of General Semantics. Washington: Catholic Univ. of America Press, 1958. Pp. 195.

Greenidge, C. W. W. Slavery. New York: Macmillan, 1958. Pp. 235.

Griffin, D. R. Listening in the Dark. New Haven: Yale Univ. Press, 1958. Pp. 413.

Gruenberg, S. M. The Parents' Guide to Everyday Problems of Boys and Girls. New York: Random House, 1958. Pp. 363.

Hammer, E. P., *Ed.* The Clinical Application of Projective Drawings. Springfield: Thomas, 1958. Pp. 663.

Harper, R. S. Introductory Psychology. Boston: Allyn & Bacon, 1958. Pp. 389.

Hartmann, H. Ego Psychology and the Problem of Adaptation. New York: International Univ. Press, 1958. Pp. 121.

Hauser, P. M., *Ed.* Population and World Politics. Glencoe: Free Press, 1958. Pp. 297.

Haydu, G. G. The Architecture of Sanity. New York: Julian, 1958. Pp. 238.

Hebb, D. O. A Textbook of Psychology. Philadelphia: Saunders, 1958. Pp. 276.

Heider, F. The Psychology of Interpersonal Relations. New York: Wiley, 1958. Pp. 322.

Hendrick, I. Facts and Theories of Psychoanalysis. New York: Knopf, 1958. Pp. 385.

Hoey, A. F. A Comparative Study of the Problems and Guidance Resources of Catholic College Women. Washington: Catholic Univ. Press, 1957. Pp. 220.

Hogben, L. Statistical Theory. New York: Norton, 1958. Pp. 510.

Holt, R. R., & Luborsky, L. Personality Patterns of Psychiatrists. New York: Basic Books, 1958. Pp. 386.

Hook, S., *Ed.* Determinism and Freedom in the Age of Modern Science. New York: New York Univ. Press, 1958. Pp. 237.

Hughes, E. C., Hughes, H. N., & Deutscher, I. Twenty Thousand Nurses Tell Their Story. Philadelphia: Lippincott, 1958. Pp. 280.

Inhelder, B., & Piaget, J. The Growth of Logical Thinking from Childhood to Adolescence. New York: Basic Books, 1958. Pp. 356.

Jahoda, M. Current Concepts of Positive Mental Health. New York: Basic Books, 1958. Pp. 136.

Janis, I. L. Psychological Stress. New York: Wiley, 1958. Pp. 439.

Karon, B. P. The Negro Personality. New York: Springer, 1958. Pp. 184.

Katzenelbogen, S. Analyzing Psychotherapy. New York: Philosophical, 1958. Pp. 126.

Kline, M. V. Freud and Hypnosis. New York: Julian, 1958. Pp. 207.

Konopka, G. Edward C. Lindeman and Social Work Philosophy. Minneapolis: Univ. Minnesota Press, 1958. Pp. 220.

Kramer, E. Art Therapy in a Children's Community. Springfield: Thomas, 1958. Pp. 238.

Krugman, M., *Ed.* Orthopsychiatry and the School. New York: American Orthopsychiatric Association, 1958. Pp. 265.

Lancaster, E., with Poling, J. The Final Fact of Eve. New York: McGraw-Hill, 1958. Pp. 290.

Lawson, C. A. Language, Thought, and the Human Mind. East Lansing: Michigan State Univ. Press, 1958. Pp. 117.

Leavitt, H. J. Managerial Psychology. Chicago: Univ. Chicago Press, 1958. Pp. 335.

Lockerby, F. K. Communication for Nurses. St. Louis: Mosby, 1958. Pp. 175.

Lorand, S., & Balint, M., *Eds.* Perversions: Psychodynamics and Therapy. New York: Random House, 1956. Pp. 307.

Luszki, M. B. Interdisciplinary Team Research. New York: New York Univ. Press, 1958. Pp. 355.

Maccoby, E. E., Newcomb, T. M., & Hartley, E. L., *Eds.* Readings in Social Psychology. New York: Holt, 1958. Pp. 674.

MacIntyre, A. C. The Unconscious: A Conceptual Analysis. New York: Humanities, 1958. Pp. 100.

Maier, N. R. F. The Appraisal Interview. New York: Wiley, 1958. Pp. 246.

Masserman, J. H. Science and Psychoanalysis. Vol. 1. New York: Grune & Stratton, 1958. Pp. 201.

Masserman, J. H., & Moreno, J. L., *Eds.* Progress in Psychotherapy. Vol. III. New York: Grune & Stratton, 1958. Pp. 324.

May, R., Angel, E., & Ellenberger, H. F., *Eds.* Existence: A New Dimension in Psychiatry and Psychology. New York: Basic Books, 1958. Pp. 445.

McClelland, D. C., Baldwin, A. L., Bronfenbrenner, U., & Strodtbeck, F. L. Talent and Society. Princeton: Van Nostrand, 1958. Pp. 275.

McCormick, T. C., & Francis, R. G. Methods of Research in the Behavioral Sciences. New York: Harper, 1958. Pp. 244.

McFarland, D. E. Management Principles and Practices. New York: Macmillan, 1958. Pp. 612.

McGuigan, F. J., & Calvin, A. D. Current Studies in Psychology. New York: Appleton-Century-Crofts, 1958. Pp. 226.

McKinney, F. Counseling for Personal Adjustment in Schools and Colleges. Boston: Houghton Mifflin, 1958. Pp. 584.

Mead, M., Michael, D. N., Lasswell, H. D., & Frank, L. K. Man in Space. New York: New York Academy of Sciences, 1958. Pp. 165-214.

Melwik, C., & Leites, N. The House Without Windows. Evanston: Row, Peterson, 1958. Pp. 358.

Menninger, K. Theory of Psychoanalytic Technique. New York: Basic Books, 1958. Pp. 206.

Mickel, J. Human Communication and General Semantics. New York: New Voices, 1958. Pp. 102.

Miller, D. R., & Swanson, G. E. The Changing American Parent. New York: Wiley, 1958. Pp. 302.

Moll, A. The Study of Hypnosis. New York: Julian, 1958. Pp. 410.
Money, J. The Psychologic Study of Man. Springfield: Thomas, 1957. Pp. 216.
Moore, H. E. Tornadoes Over Texas. Austin: Univ. Texas Press, 1958. Pp. 334.
Mudd, E. H., Stone, A., Karpf, M. J., & Nelson, J. F., *Eds.* Marriage Counseling. New York: Association Press, 1958. Pp. 488.
Murphy, G. Human Potentialities. New York: Basic Books, 1958. Pp. 340.
Nash, M. Machine Age Maya. Glencoe: Free Press, 1958. Pp. 118.
Nice, R. W., *Ed.* Crime and Insanity. New York: Philosophical, 1958. Pp. 280.
Parad, H. J., *Ed.* Ego Psychology and Dynamic Casework. New York: Family Service Association, 1958. Pp. 282.
Pease, O. The Responsibilities of American Advertising. New Haven: Yale Univ. Press, 1958. Pp. 232.
Peters, R. S. The Concept of Motivation. New York: Humanities, 1958. Pp. 166.
Phelps, H. S., & Henderson. Population in its Human Aspects. New York: Appleton-Century-Crofts, 1958. Pp. 512.
Reik, T. The Search Within. New York: Grove, 1958. Pp. 657.
Rinkel, M., & Denber, H. C. B., *Eds.* Chemical Concepts of Psychosis. New York: McDowell, Obelensky, 1958. Pp. 485.
Rothney, J. W. M. Guidance Practices and Results. New York: Harper, 1958. Pp. 542.
Roucek, J. S., *Ed.* Juvenile Delinquency. New York: Philosophical, 1958. Pp. 370.
Sargent, S. S., & Williamson, R. C. Social Psychology. New York: Ronald Press, 1958. Pp. 649.
Sartain, A. Q., North, A. J., Strange, J. R., & Chapman, H. M. Psychology: Understanding Human Behavior. New York: McGraw-Hill, 1958. Pp. 444.
Sayles, L. R. Behavior of Industrial Work Groups. New York: Wiley, 1958. Pp. 182.
Seward, G. S., & Seward, J. P. Current Psychological Issues. New York: Henry Holt, 1958. Pp. 360.
Torgerson, W. S. Theory and Methods of Scaling. New York: Wiley, 1958. Pp. 460.
Woodworth, H. Sanity, Unheard Of. Victoria, B.C.: Sumas Ltd., 1958. Pp. 109.

[illegible] PSYCHOLOGICAL [illegible]